Title	Macroprudential policy and forecasting using Hybrid DSGE models with financial frictions and State space Markov-Switching TVP-VARs
Authors(s)	Bekiros, Stelios D., Paccagnini, Alessia
Publication date	2015-10
Publication information	Bekiros, Stelios D., and Alessia Paccagnini. "Macroprudential Policy and Forecasting Using Hybrid DSGE Models with Financial Frictions and State Space Markov-Switching TVP-VARs." Cambridge University Press, October 2015. https://doi.org/10.1017/S1365100513000953.
Publisher	Cambridge University Press
Item record/more information	http://hdl.handle.net/10197/7333
Publisher's version (DOI)	10.1017/S1365100513000953

Macroprudential policy and forecasting using Hybrid DSGE models with financial frictions and State space Markov-Switching TVP-VARs*

Stelios D. Bekiros[a,b,c†]

[a]European University Institute (EUI), Florence, Italy

[b]Rimini Centre for Economic Analysis (RCEA), Rimini, Italy

[c]Athens University of Economics & Business (AUEB), Athens, Greece

Alessia Paccagnini[‡]

Università degli Studi di Milano-Bicocca, Milano, Italy

Abstract

The financial crisis has shown that neither VAR nor simple DSGE macro-modelling is fully ready to deal explicitly with the issue of financial frictions and stability that play an important role in modern business cycle fluctuations. DSGE models that incorporate financial frictions could deal with the transmission mechanism of standard shocks changes and how monetary policy is affected by the presence of frictions, as well as with optimal macroprudential policies and the impact of capital requirements. However, DSGE and VAR models are still linear and they do not consider time-variation in parameters that could account for inherent nonlinearities and capture the adaptive underlying structure of the economy in a robust manner. In this paper, a novel multivariate state-space estimation method for time-varying VAR models is introduced. As an extension to the standard homoskedastic TVP-VAR we employ a Markov-switching heteroskedastic error structure. Overall, we conduct a comparative empirical analysis of the out-of-sample predictive performance of simple and hybrid DSGE models against standard VARs, Bayesian VARs, Factor Augmented VARs and TVP-VARs, using datasets from the US economy. We apply advanced Bayesian and Quasi-optimal filtering techniques in estimating and forecasting the models. From a modelling point of view we focus on the interaction of frictions both at firms' level and in the banking sector in order to examine the transmission mechanism of the shocks and to reflect on the response of the monetary policy to increases in interest rate spreads, especially after the financial crisis. A first attempt is made to find macro-financial micro-founded DSGE models as well as adaptive TVP-VARs, which are able to reflect reality better.

JEL Classification: C32, C11, C15

Keywords: Forecasting, Marginal data density, Financial frictions, DSGE models, Time-varying coefficients, Quasi-optimal filtering

---

*The first author is supported by the Marie Curie Fellowship (FP7-PEOPLE-2011-CIG, N° 303854) under the 7th European Community Framework Programme. The second author acknowledges financial support from "Dote ricercatori": FSE, Regione Lombardia. The usual disclaimers apply.

†Corresponding author; Department of Economics, Via della Piazzuola 43, I-50133, Florence, Italy; Tel.: +39 055 4685916; fax: +39 055 4685 902.
Rimini Centre for Economic Analysis (RCEA), Via Patara 3, 47900, Rimini, Italy; Tel/: +39 0541 434 142; fax: +39 0541 55 431.
Department of Acc. & Finance, 76 Patission str., GR10434, Athens, Greece; Tel./fax: +30 210 8203453.
E-mail addresses: stelios.bekiros@eui.eu , bekiros@aueb.gr

‡Department of Economics, Piazza Ateneo Nuovo 1, 20126, Milano, Italy; Tel.: +39 026 4483 046; fax: +39 026 4483 085;
E-mail address: alessia.paccagnini@unimib.it

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models are widely used among central banks and institutions as well as in the academia in order to test alternative monetary, fiscal and macroprudential policies under different scenarios. However, the financial crisis shows that the workhorse of contemporary DSGE modelling is not fully ready to deal explicitly with the issue of credit frictions and financial stability, which play an important role in modern business cycle. In recent times there have been some efforts to model financial factors: Goodfriend and McCallum (2007), Christiano *et al.* (2010), Curdía and Woodford (2010), Gerali *et al.* (2010), Gertler and Kiyotaki (2010) and Gertler and Karadi (2011) constitute some promising attempts to study the effects of financial intermediation on business cycle fluctuations and policy design. In particular, DSGE models with financial frictions could deal with many issues, such as: (i) the effect of financial shocks on real variables, (ii) how the transmission mechanism of “standard shocks” changes in presence of imperfect capital markets (iii) how optimal monetary policy is affected by the presence of financial frictions and (iv) optimal macroprudential policies with respect to the impact on capital requirements. In the literature there are three main strands of research on financial frictions: (i) financial frictions at the level of firms, following the seminal contribution of Bernanke *et al.* (1999), (ii) frictions at the level of households in the form of collateral constraint introduced by Kiyotaki and Moore (1997) and Iacoviello (2005) and (iii) frictions at the level of financial intermediaries as in Gertler and Karadi (2011). These different modelling techniques capture the problem of asymmetric information between borrowers and lenders, which leads to a constraint in the borrowers’ ability to obtain credit.

The DSGE models appear to be particularly suited for evaluating the consequences of alternative macroeconomic policies, as shown in the works of Smets and Wouters (2003, 2004), Del Negro and Schorfheide (2004), Adolfson *et al.* (2008) and Christiano *et al.* (2005). However, the calibrated dynamic stochastic general equilibrium models face many important challenges such as the fragility of parameter estimates, statistical fit and the weak reliability of policy predictions as reported in Stock and Watson (2001), Ireland (2004) and Schorfheide (2010). In recent years Bayesian estimation of DSGE models has become popular for many reasons, mainly because it is a system-based estimation approach that offers the advantage of incorporating assumptions about the parameters based on economic theory. Recently, increasing efforts have been undertaken to use DSGE models for forecasting. DSGE models were not considered as forecasting tools until the works of Smets and Wouters (2003, 2004) on the predictability of DSGE models compared to alternative non-structural models. In the macro-econometric literature, hybrid or mixture DSGE models have become popular for dealing with some of the model misspecifications as well as the trade-off between theoretical coherence and empirical fit (Schorfheide, 2010). They are categorized in additive hybrid models and hierarchical hybrid models. The hybrid models provide a complete analysis of the data law of motion and better capture the dynamic properties of the DSGE models. In the recent literature, different attempts of hybrid models have been introduced for solving, estimating and forecasting with DSGEs. Sargent (1989) and Altug (1989) proposed augmenting a DSGE model with measurement error terms that follow a first order autoregressive process, known as the DSGE-AR approach. Ireland (2004) proposed a method that is similar to the DSGE-AR, but imposing no restriction on the measurement errors, assuming that residuals follow a first-order vector autoregression (DSGE-AR à l’Ireland). A different approach called DSGE-VAR was proposed by Del Negro and Schorfheide (2004) and was based on the works DeJong *et al.* (1996) and Ingram and Whiteman (1994). The main idea behind the DSGE-VAR is the use of the VAR representation as an econometric tool for empirical validation, combining prior information derived from the DSGE model in estimation. However, it has several problems. One of the main problems in finding a statistical representation for the data by using a VAR, is overfitting due to the inclusion of too many lags and too many variables, some of which may be insignificant. The problem of overfitting results in multicollinearity and loss of degrees of freedom, leading to inefficient estimates and large out-of-sample forecasting errors. It is possible to overcome this problem by using the well-known "Minnesota" priors (Doan *et al.*, 1984). The use of "Minnesota" priors has been proposed to shrink the parameters space and thus overcome the curse of dimensionality. Following this idea in combining the DSGE model information and the VAR representation, two alternative econometric tools have been also introduced: the DSGE-FAVAR (Consolo *et al.*, 2009) and the Augmented VAR-DSGE model (Fernández-de-Córdoba and Torres, 2010). The main idea behind the Factor Augmented DSGE (DSGE-FAVAR) is the use of factors to improve the statistical identification in validating the models. Consequently, the VAR representation is replaced by a FAVAR model as the statistical benchmark.

Nevertheless, DSGE and VAR modelling can be subject to the Lucas critique (Lucas, 1976) in that atheoretic statistical models as well as micro-founded models fail to take into account inherent nonlinearities of the economy. Even though VAR models are proven to be reliable tools for data description and forecasting, they can only be utilized in the analysis of stationary macro-series, and in many cases stationarity assumptions are too restrictive. In these cases, the use of time-varying parameters seems to be an attractive alternative in order to capture nonlinear economic relationships. The time-varying properties are very useful, because they relax stationarity assumptions, and they also provide a simple interpretation of functional coefficients. Time varying autoregression (TVP-VAR) models have been developed since the early 1980's. Prado et al. (2001) offers an excellent review. Primiceri (2005) used them extensively in analyzing macroeconomic policy issues. The TVP-VAR model enables capturing a possible time-varying nature of underlying structure in the economy in a flexible and robust manner. TVP-VAR models led to new methods of time series decomposition and analysis as presented with applications in Primiceri (2005). In this paper, a novel time-varying multivariate state-space estimation method for TVP-VAR processes is proposed both for homoskedastic and heteroskedastic error structures. As an alternative to the homoskedastic TVP-VAR we assume that the error structure of the state space Kalman filter is dependent on state variables, which are unobserved discrete-time, discrete-state Markov process, thus providing a Markov-switching heteroskedasticity. While a simple Markov-switching variance model fails to incorporate the learning process of agents, the classic TPV-VAR model fails to incorporate uncertainty that changes due to future asymmetric random shocks. In this work we consider a more general model in which both types of uncertainty are incorporated. For the TVP-VAR models, the parameters are estimated using a multivariate specification of the standard Kalman set-up (Harvey, 1990) with extended quasi-optimal filtering in particular for the TVP-VAR with Markov-switching heteroskedasticity. The likelihood estimation of the TVP-VAR is performed with a suitable multivariate extension of Kim (1993) and Kim and Nelson (1999a, 1999b) method.

Overall, in this work, we conduct an exhaustive empirical exercise that includes the comparison of the out-of-sample predictive performance of estimated simple and hybrid DSGE models with that of standard VARs, Bayesian VARs and Factor Augmented VARs as well as DSGE-VAR and Factor Augmented DSGE models estimated on the same data set for the US economy. In addition, we assess comparatively the forecastability of two time-varying parameter autoregressive models (TVP-VAR) models with homoskedastic and heteroskedastic errors against all of the aforementioned model specifications, in an attempt to investigate inherent nonlinearities of the economy that cannot be captured by the VAR and DSGE class models. The DSGE model is obtained by augmenting the small-scale model with financial frictions, following Christiano et al. (2009) and Del Negro and Schorfheide (2012). We start considering the Smets-Wouters model with financial frictions proposed by Del Negro and Schorfheide (2012). After the recent financial crisis, the spread is a key variable. As in Christiano, Motto and Rostagno (2009) we measure the spread as the annualized Moody's Seasoned BAA corporate bond yield spread over the 10-Year Treasury note yield at constant maturity, and we include it to reflect on the response of the monetary policy to increases in interest rate spreads. The variety of estimation techniques employed makes it possible to compare estimations according to different criteria: analyzing the posterior means of key parameters, calculating the marginal data density, building Bayes Factors and forecasting comparison. The main purpose of this paper is to compare different econometrics strategies in evaluating a DSGE economy, but mainly to stress the importance of considering financial variables in particular for the US economy during and after the recent financial crisis, and their incorporation in DSGE and TVP-VAR models. We focus on many different specifications of the DSGE models, i.e., the simple DSGE, the DSGE-VAR and specifically on the Factor Augmented DSGE (DSGE-FAVAR) model with emphasis on Bayesian estimation, as well as on two specifications for TVP-VAR models. We use time series data from 1960:Q4 to 2010:Q4 for the real GDP, the harmonized Consumer Price Index, the nominal short-term federal funds interest rate and the yield spread, and we produce their forecasts for the out-of-sample testing period 1997:Q1-2010:Q4. The motivation comes from a group of recent papers that compares the forecasting performance of DSGE against VAR models, e.g., Smets and Wouters (2004), Ireland (2004), Del Negro and Schorfheide (2004), Del Negro et al. (2007), Adolfson et al. (2008), among others. The remainder of this paper is organized as follows. Section 2 describes the standard and Bayesian VAR as well as the Factor Augmented VAR model. In section 3 the proposed DSGE model with financial frictions is analyzed, and the hybrid DSGE-VAR and DSGE-FAVAR models are described in detail. Section 4 presents the time-varying multivariate state-space homoscedastic TVP-VAR model as well as a Markov-switching heteroskedastic set-up. In section 5 the data are described and the empirical results of the comparative forecasting evaluation are illustrated and analyzed. Finally, section 6 concludes.

2 Vector Autoregressive Models

The standard unrestricted VAR as suggested by Sims (1980) has the following compact format

$$\mathbf{Y}_t = \mathbf{X}_t \mathbf{\Phi} + \mathbf{U} \tag{1}$$

where $\mathbf{Y}_t$ is a $(T \times n)$ matrix with rows Y_t', and $\mathbf{X}$ is a $(T \times k)$ matrix $(k = 1 + np, p =$number of lags) with rows $X_t' = [1, Y_{t-1}', ..., Y_{t-p}']$. $\mathbf{U}$ is a $(T \times n)$ matrix with rows u_t', $\mathbf{\Phi}$ is a $(k \times n) = [\Phi_{0,} \Phi_1, ..., \Phi_p]'$, while the one-step ahead forecast errors u_t have a multivariate $N(0, \Sigma_u)$ conditional on past observations of Y.

2.1 Bayesian VAR

One of main problems in using VAR models is that many parameters need to be estimated, although some of them may be insignificant. Instead of eliminating longer lags, the BVAR imposes restrictions on these coefficients by assuming that they are more likely to be near zero than the coefficients on shorter lags (Litterman, 1981; Doan *et al.,* 1984; Todd, 1984; Litterman, 1986; Spencer, 1993). Obviously, if there are strong effects from less important variables, the data can counter this assumption. Usually, the restrictions are imposed by specifying normal prior distributions with zero means and small standard deviations for all coefficients, with a decreasing standard deviation as the lags increase. The only exception is the coefficient on a variable's first lag that has a mean of unity. Litterman (1981) used a diffuse prior for the constant. The means of the prior are popularly called the "Minnesota Priors" due to the development of the idea at the University of Minnesota and the Federal Reserve Bank at Minneapolis[1]. Formally speaking, these prior means can be written as follows

$$\Phi_i \sim N(1, \sigma^2_{\Phi_i}) \text{ and } \Phi_j \sim N(0, \sigma^2_{\Phi_j}), \tag{2}$$

where Φ_i denotes the coefficients associated with the lagged dependent variables in each equation of the VAR, while Φ_j represents any other coefficient. The prior variances $\sigma^2_{\Phi_i}$ and $\sigma^2_{\Phi_j}$ specify the uncertainty of the prior means, $\Phi_i = 1$ and $\Phi_j = 0$, respectively. In this study, we impose their prior mean on the first own lag for variables in growth rate, such as a white noise setting (Del Negro and Schorfheide, 2004; Adolfson *et al.,* 2007; Banbura *et al.,* 2010). Instead, for level variables, we use the classical Minnesota prior (Del Negro and Schorfheide, 2004). The specification of the standard deviation of the distribution of the prior imposed on variable j in equation i at lag m, for all i, j and m, denoted by $S(i, j, m)$, is specified as follows

$$S(i, j, m) = [w \times g(m) \times F(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j}, \tag{3}$$

where

$$F(i, j) = \begin{cases} 1 \text{ if } i = j \\ k_{ij} \text{ otherwise, } 0 \leq k_{ij} \leq 1 \end{cases} \tag{4}$$

is the tightness of variable j in equation i relative to variable i and by increasing the interaction, i.e. it is possible for the value of k_{ij} to loosen the prior (Dua and Ray, 1995). The ratio $\frac{\hat{\sigma}_i}{\hat{\sigma}_j}$ consists of estimated standard errors of the univariate autoregression, for variables i and j. This ratio scales the variables to account for differences in the units of measurement, without taking into account the magnitudes of the variables. The term w measures the standard deviation on the first lag, and also indicates the overall tightness. A decrease in the value of w results in a tighter prior. The function $g(m) = m^{-d}$, $d > 0$ is the measurement of the tightness on lag m relative to lag 1, and is assumed to have a harmonic shape with a decay of d, which tightens the prior on increasing lags. Following the standard Minnesota prior settings, we choose the overall tightness (w) to be equal to 0.3, while the lag decay (d) is 1 and the interaction parameter (k_{ij}) is set equal to 0.5.

2.2 Factor Augmented VAR

Recently, Stock and Watson (2002), Forni and Reichlin (1996, 1998) and Forni *et al.* (1999, 2000) have shown that very large macroeconomic datasets can be properly modelled using dynamic factor models, where the factors can be considered as an "exhaustive summary of the information" in the data. The rationale underlying dynamic factor models is that the behavior of several variables is driven by few common forces, the factors, plus idiosyncratic shocks. Hence, the factors-approach can be useful in alleviating the omitted variable problem in empirical analysis using traditional small-scale models. Bernanke and Boivin (2003) and Bernanke *et al.* (2005) utilized factors in the estimation of VAR to generate a more general specification. Chudik and Pesaran (2011) illustrated how a VAR augmented by factors could help in keeping the number of estimated parameters under control without loosing relevant information.

[1] The basic principle behind the "Minnesota" prior is that all equations are centered around a random walk with drift. This idea has been modified by Kadiyala and Karlsson (1997) and Sims and Zha (1998). In Ingram and Whiteman (1994), a real business cycle model is used to generate a prior for a reduced form VAR, as a development of the "Minnesota" priors procedure. Also, a prior is placed on the parameters of a simple linearized DSGE, which is then compared with a Bayesian VAR in a forecasting exercise. Smets and Wouters (2003) extend this to medium scale New Keynesian models used in policy analysis. This approach has the advantage of providing information about which behavioural mechanisms produce forecast error or policy scenarios. However, it seems that it often fails to empirically fit compared to models with no behavioural structure. In Del Negro and Schorfheide (2004) and Del Negro *et al.* (2007), a DSGE prior is also developed for a VAR.

Let $\mathbf{X}_t$ denote an $N \times 1$ vector of economic time series and $\mathbf{Y}_t$ a vector of $M \times 1$ observable macroeconomic variables which are a subset of $\mathbf{X}_t$. In this context, most of the information contained in $\mathbf{X}_t$ is captured by $\mathbf{F}_t$, a $k \times 1$ vector of unobserved factors. The factors are interpreted as an addition to the observed variables, as common forces driving the dynamics of the economy. The relation between the "informational" time series $\mathbf{X}_t$, the observed variables $\mathbf{Y}_t$ and the factors $\mathbf{F}_t$ is represented by the following dynamic factor model:

$$\mathbf{X}_t = \mathbf{\Lambda}^f \mathbf{F} + \mathbf{\Lambda}^y \mathbf{Y}_t + e_t \tag{5}$$

where $\mathbf{\Lambda}^f$ is a $N \times k$ matrix of factor loadings, $\mathbf{\Lambda}^y$ is a $N \times M$ matrix of coefficients that bridge the observable $\mathbf{Y}_t$ and the macroeconomic dataset, and e_t is the vector of $N \times 1$ error terms. These terms are mean zero, normal distributed, and uncorrelated with a small cross-correlation. In fact, the estimator allows for some cross-correlation in e_t that must vanish as N goes to infinity. This representation nests also models where $\mathbf{X}_t$ depends on lagged values of the factors (Stock and Watson, 2002). For the estimation of the FAVAR model equation (5), we follow the two-step principal components approach proposed by Bernanke *et al.* (2005). In the first step factors are obtained from the observation equation by imposing the orthogonality restriction $\mathbf{F}'\mathbf{F}/T = \mathbf{I}$.This implies that $\hat{\mathbf{F}} = \sqrt{T}\hat{\mathbf{G}}$, where $\hat{\mathbf{G}}$ are the eigenvectors corresponding to the K largest eigenvalues of $\mathbf{X}\mathbf{X}^{'}$, sorted in descending order. Stock and Watson (2002) showed that the factors can be consistently estimated by the first r principal components of $\mathbf{X}$, even in the presence of moderate changes in the loading matrix $\mathbf{\Lambda}$. For this result to hold it is important that the estimated number of factors, k, is larger or equal than the true number r. Bai and Ng (2000) proposed a set of selection criteria to choose k that are generalizations of the BIC and AIC criteria. In the second step, we estimate the FAVAR equation replacing $\mathbf{F}_t$ by $\hat{\mathbf{F}}_t$. Following Bernanke *et al.* (2005), $\mathbf{Y}_t$ is removed from the space covered by the principal components. In a recent paper, Boivin *et al.* (2009) impose the constraint that $\mathbf{Y}_t$ is one of the common components in the first step, guaranteeing that the estimated latent factors $\hat{\mathbf{F}}_t$ recover the common dynamics which are not captured by $\mathbf{Y}_t$. The authors, comparing the two methodologies, concluded that the results are similar. As in Bernanke *et al.* (2005) we partition the matrix $\mathbf{X}_t$ in two categories of information variables: slow-moving and fast-moving. Slow-moving variables (e.g., real variables such as wages or spending) do not respond contemporaneously to unanticipated changes in monetary policy, while fast-moving (e.g., interest rates) respond contemporaneously to monetary shocks. We proceed to extracting two factors from slow variables and one factor from fast variables and we call them respectively "slow factors" and "fast factor". As suggested by Bai and Ng (2000) information criteria can be used to determine the number of factors but, as they are not so decisive, one can limit the number of factors to three (two slows and one fast) to strike a balance between the variance of the original series explained by the principal components and the difference in the parameterization of the VAR and the FAVAR. It is also worth noting that the factors are not uniquely identified, but this is not a problem in our context because we will not attempt a structural interpretation of the estimated factors. Finally, having determined the number of factors, we specify a Factor Augmented VAR by considering only one-lag of the factors according to BIC criterion. The potential identification of the macroeconomic shocks can be performed according to Bernanke *et al.* (2005) using the Cholesky decomposition.

3 DSGE model with financial frictions

The model proposed is a simple DSGE model obtained as a special case of the Smets and Wouters (2007) model. We augment the small scale DSGE model by financial frictions as shown in Del Negro and Schorfheide (2012), following the work of Bernanke *et al.* (1999) and Christiano *et al.* (2009). We maintain a simplified version of capital to introduce the concept of financial friction with a small scale DSGE model. We follow the approach proposed by Del Negro and Schorfheide (2012), in that we detrend the non-stationary model variables by a stochastic trend rather than a determinist trend. Del Negro and Schorfheide (2012) propose this approach to express almost all equilibrium conditions in a way that encompasses both the trend-stationary total factor productivity process in Smets and Wouters (2007) as well as when the technology follows a unit root process.

Let $\tilde{z}_t$ be the linearly detrended log productivity process which follows the autoregressive AR(1) law of motion as exogenous shock:

$$\tilde{z}_t = \rho_z \tilde{z}_{t-1} + \sigma_z \epsilon_{z,t}, \tag{6}$$

where $\epsilon_{z,t}$ is iid standard normal. Equation (6) implies that the growth rate of the trend process evolves according to:

$$z_t = \ln(\frac{Z_t}{Z_{t-1}}) - \gamma = \frac{1}{1-\alpha}(\rho_z - 1)\tilde{z}_{t-1} + \frac{1}{1-\alpha}\sigma_z \epsilon_{z,t}$$

since we detrend all non stationary variables by $Z_t = e^{\gamma t + \frac{1}{1-\alpha}\widetilde{z}_t}$, where γ is the steady state growth rate of the economy. The consumption Euler equation takes the form:

$$c_t = E_t[c_{t+1} + z_{t+1}] - \frac{1}{\sigma_c}(R_t - E_t[\pi_{t+1}]) \tag{7}$$

where c_t is consumption, R_t is the nominal interest rate, $R_t - E_t[\pi_{t+1}]$ is the riskless return, π_t is inflation and σ_c is the relative degree of risk aversion. In this model we do not assume investment, consequently the capital accumulation process is given:

$$\overline{k}_t = \overline{k}_{t-1} - z_t + \rho_\mu \mu_{t-1} + \sigma_\mu \epsilon_{\mu,t}$$

The arbitrage condition between the return to capital and the riskless rate is modified as proposed by Del Negro and Schorfheide (2012):

$$E[\widetilde{R}^k_{t+1} - R_t] = \zeta_{sp}(\overline{k}_t - n_t) + \widetilde{\sigma}_{\omega,t}$$

and

$$\widetilde{R}^k_t - \pi_t = \frac{r^k_*}{r^k_* + (1-\delta)} r^k_* \tag{8}$$

where r^k_* is the rental rate of capital of steady state, δ is the depreciation rate, $\widetilde{R}^k_t$ is the gross nominal return on capital for entrepreneurs, n_t is entrepreneurial equity which depends on equation (8) and $\widetilde{\sigma}_{\omega,t}$ captures mean-preserving changes in the cross-sectional dispersion of ability across entrepreneurs (Christiano *et al.*, 2009) and follows an AR(1) process with parameters $\rho_{\sigma\omega}$ and $\sigma_{\sigma\omega}$. The first condition determines the spread between the expected return on capital and the riskless rate (if $\zeta_{sp} = 0$, the financial friction shocks are zero), while the second condition defines the return on capital. Capital is subject to variable capacity utilization u_t. The relationship between $\overline{k}_t$ and the amount of capital effectively rented out to firms k_t is:

$$k_t = u_t - z_t + \overline{k}_{t-1}$$

The optimality condition determining the rate of utilization is given by:

$$\frac{1-\psi}{\psi} r^k_t = u_t$$

where ψ captures the utilization costs in terms of foregone consumption. From the optimality conditions of goods producers it follows that all firms have the same capital-labor ratio:

$$k_t = w_t - r^k_t + L_t$$

The real marginal costs are given by:

$$mc_t = w_t + \alpha L_t - \alpha k_t$$

where α is the income share of capital in the production.

The Phillips curve takes the form:

$$\pi_t = \frac{(1-\zeta_p\beta)(1-\zeta_p)}{(1+\iota_p\beta)\zeta_p} mc_t + \frac{\beta}{1+\iota_p\beta} E_t[\pi_{t+1}] + \frac{\iota_p}{1+\iota_p\beta}\pi_{t-1} \tag{9}$$

where ζ_p, and ι_p are the Calvo parameter and the degree of indexation. We assume that the central bank only reacts to inflation and output growth and that the monetary policy shock is iid. The policy rule is:

$$R_t = \rho_R R_{t-1} + (1-\rho_R)\left[\psi_1 \pi_t + \psi_2(y_t - y_{t-1} + z_t)\right] + \sigma_R \epsilon_{R,t} \tag{10}$$

where $\epsilon_{R,t}$ is iid standard normal.

The production function is:

$$y_t = \alpha k_t + (1-\alpha) L_t$$

The aggregate resource constraint is:

$$y_t = c_t + g_t + k_t$$

where g_t is an exogenous shock process, evolving such as an AR(1):

$$g_t = \rho_g g_{t-1} + \sigma_g \epsilon_{g,t}$$

and $\epsilon_{g,t}$ is iid standard normal.

The previous set of equations can be recasted into a set of matrices $(\Gamma_0, \Gamma_1, C, \Psi, \Pi)$ accordingly to the definition of the vectors $\tilde{Z}_t$ and ϵ_t

$$\Gamma_0 \tilde{Z}_t = C + \Gamma_1 \tilde{Z}_{t-1} + \Psi \epsilon_t + \Pi \eta_t \tag{11}$$

where η_{t+1}, such that $E_t \eta_{t+1} \equiv E_t (y_{t+1} - E_t y_{t+1}) = 0$, is the expectations error. As a solution to (11), we obtain the following transition equation as a policy function

$$\tilde{Z}_t = T(\theta) \tilde{Z}_{t-1} + R(\theta) \epsilon_t \tag{12}$$

and in order to provide the mapping between the observable data and those computed as deviations from the steady state of the model. The small-scale model is estimated based on four quarterly macroeconomic time series: The measurement equations for real output growth, inflation, short term interest rate and spread:

$$\begin{aligned} \Delta \ln y_t &= \ln \gamma + \Delta y_t + z_t \\ \Delta \ln P_t &= \ln \pi^* + \pi_t \\ \ln R_t^a &= 4 \left[(\ln R^* + \ln \pi^*) + R_t \right] \\ SP_t &= SP_* + 100 E_t [\widetilde{R}_{t+1}^k - R_t] \end{aligned} \tag{13}$$

where all variables are measured in percent and π_*, R_* and SP_* measure the steady state level of inflation, short term interest rate and spread. This can be also casted into matrices as

$$Y_t = \Lambda_0(\theta) + \Lambda_1(\theta) \tilde{Z}_t + v_t \tag{14}$$

where $Y_t = (\Delta \ln y_t, \Delta \ln P_t, \ln R_t, SP_t)'$, $v_t = 0$ and Λ_0 and Λ_1 are defined accordingly. For completeness, we write the matrices T, R, Λ_0 and Λ_1 as a function of the structural parameters in the model, θ. Such a formulation derives from the rational expectations solution. The evolution of the variables of interest, Y_t, is therefore determined by (12) and (14) which impose a set of restrictions across the parameters on the moving average (MA) representation. Given that the MA representation can be very closely approximated by a finite order VAR representation, Del Negro and Schorfheide (2004) propose to evaluate the DSGE model by assessing the validity of the restrictions imposed by such a model with respect to an unrestricted VAR representation. The choice of the variables to be included in the VAR is however completely driven by those entering in the DSGE model regardless of the statistical goodness of the unrestricted VAR.

3.1 Estimation of the linearized DSGE Model

Several econometric procedures have been proposed to parameterize and evaluate DSGE models. Kydland and Prescott (1982) use calibration, Christiano and Eichenbaum (1992) consider the generalized method of moments (GMM) estimation of equilibrium relationships, while Rotemberg and Woodford (1997) and Christiano *et al.* (2005) use the minimum distance estimation based on the discrepancy among VAR and DSGE model impulse response functions. Moreover the full-information likelihood-based estimation is considered by Altug (1989), McGrattan (1994), Leeper and Sims (1994) and Kim (2000). In last years, Bayesian estimation became very popular. According to An and Schorfheide (2007) there are essentially three main characteristics. Firstly, the Bayesian estimation is system-based and fits the solved DSGE model to a vector of aggregate time series, as opposed to the GMM which is based on equilibrium relationships, such as the Euler equation for the consumption or the monetary policy rule. Secondly, it is based on the likelihood function generated by the DSGE model rather than the discrepancy between DSGE responses and VAR impulse responses. Thirdly, prior distributions can be used to incorporate additional information into the parameter estimation.

Priors distributions are important to estimate DSGE models. According to An and Schorfheide (2007) priors might downweigh regions of the parameter space that are at odds with observations which are not contained in the estimation sample. Priors could add curvature to a likelihood function that is (nearly) flat for some parameters, given

Table 1: Prior Distributions for the DSGE model parameters

Name	Density	Mean	Standard deviation
		Policy Parameters	
ψ_1	Gamma	1.500	0.250
ψ_2	Gamma	0.120	0.050
ρ_R	Beta	0.750	0.100
		Normal Rigidities Parameters	
ς_p	Beta	0.500	0.100
		Other Endogenous Propagation and Steady State Parameters	
α	Normal	0.300	0.050
ι_p	Beta	0.500	0.150
r_*	Gamma	0.250	0.100
π_*	Gamma	0.620	0.100
γ	Normal	0.400	0.100
σ_c	Normal	1.500	0.370
ψ	Beta	0.50	0.150
ρ_z	Beta	0.500	0.200
ρ_μ	Beta	0.500	0.200
ρ_g	Beta	0.500	0.200
σ_Z	Inv.Gamma	0.100	2.000
σ_μ	Inv.Gamma	0.100	2.000
σ_g	Inv.Gamma	0.875	0.430
		Financial Frictions	
SP_*	Gamma	2.000	0.100
$\rho_{\sigma\omega}$	Beta	0.750	0.150
ς_{sp}	Beta	0.050	0.005
$\sigma_{\sigma\omega}$	Inv.Gamma	0.050	4.000

Note: The model parameters are fixed in Del Negro and Schorfheide (2012): $\delta = 0.025$, and $g_* = 0.18$. The Inverse Gamma priors are of the form $p(\sigma|\nu, s) \propto \sigma^{-\nu-1}e^{-\nu s^2/2\sigma^2}$, where v=4 and s equals 0.2, 0.5, and 0.7, respectively. Approximately 1.5% of the prior mass lies in the indeterminacy region of the parameter space. The prior is truncated to restrict it to the determinacy region of the DSGE model, to avoid multiple equilibria typical in rational expectations models .

a strong influence to the shape of the posterior distribution. Table 1 lists the prior distributions for the structural parameters of the DSGE model which are adopted from Del Negro and Schorfheide (2012).

In the Bayesian framework, the likelihood function is reweighted by a prior density. The prior is useful to add information which is contained in the estimation sample. Since priors are always subject to revisions, the shift from prior to posterior distribution can be considered as an indicator of the different sources of information. If the likelihood function peaks at a value that is at odds with the information that has been used to construct the prior distribution, then the marginal data density (MDD) of the DSGE model is defined as:

$$p(Y) = \int L(\theta|Y)p(\theta)d\theta$$

The marginal data density is the integral of the likelihood ($L(\theta|Y)$) taken according to the prior distribution ($p(\theta)$), that is the weighted average of likelihood where the weights are given by priors. The MDD can be used to compare different models M_i, $p(Y|M_i)$.We can rewrite the log-marginal data density as:

$$\begin{aligned} \ln(p(Y|M) &= \sum_{t=1}^{T} \ln p(y_t|Y^{t-1}, M) = \\ &= \sum_{t=1}^{T} \ln\left[\int p(y_t|Y^{t-1}, \theta, M)p(\theta|Y^{t-1}, M)d\theta\right] \end{aligned}$$

where $\ln(p(Y|M)$ can be interpreted as a predictive score (Good, 1952) and the model comparison based on posterior odds captures the relative one-step-ahead predictive performance. To compute the MDD, we consider the Geweke (1999) modified harmonic mean estimator. Harmonic mean estimators are based on the identity:

$$\frac{1}{p(Y)} = \int \frac{f(\theta)}{L(\theta|Y)p(\theta)} p(\theta|Y) d\theta$$

where $f(\theta)$ has the property that $\int f(\theta) d\theta = 1$ (Gelfand and Dey, 1994). Conditional on the choice of $f(\theta)$, an estimator is:

$$\widehat{p}_G(Y) = \left[\frac{1}{n_{sim}} \sum_{s=1}^{n_{sim}} \frac{f(\theta^{(s)})}{L(\theta^{(s)}|Y)p(\theta^{(s)})} \right]^{-1} \tag{15}$$

where $\theta^{(s)}$ is drawn from the posterior $p(\theta|Y)$. For a numerical approximation efficient, $f(\theta)$ should be chosen so that the summands are of equal magnitude. Geweke (1999) proposed to use the density of a truncated multivariate normal distribution:

$$\begin{aligned} f(\theta) &= \tau^{-1}(2\pi)^{-\frac{d}{2}} |V_\theta|^{-\frac{1}{2}} \exp\left[-0.5(\theta - \overline{\theta})' V_\theta^{-1} (\theta - \overline{\theta})\right] \\ & \times I\left\{(\theta - \overline{\theta})' V_\theta^{-1} (\theta - \overline{\theta}) \leq F_{\chi_d^2}^{-1}(\tau\tau)\right\} \end{aligned}$$

In the above $\overline{\theta}$ and V_θ are the posterior mean and covariance matrix computed from the output of the posterior simulator, d is the dimension of the parameter vector, $F_{\chi_d^2}$ is the cumulative density function of a χ^2 random variable with d degrees of freedom, and $\tau \in (0, 1)$. If the posterior of θ is in fact normal then the summands in eq. (15) are approximately constant.

3.2 The Del Negro-Schorfheide DSGE-VAR

Based on the study of Ingram and Whiteman (1994), Del Negro and Schorfheide (2004) designed the DSGE-VAR approach to improve forecasting and monetary policy analysis with VARs. Del Negro-Schorfheide's (2004) approach is to use the DSGE model to build prior distributions for the VAR. Basically, the estimation initializes with an unrestricted VAR of order p

$$\mathbf{Y}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{Y}_{t-1} + ... + \mathbf{\Phi}_p \mathbf{Y}_{t-p} + \mathbf{u}_t \tag{16}$$

In compact format:

$$\mathbf{Y} = \mathbf{X\Phi} + \mathbf{U} \tag{17}$$

$\mathbf{Y}$ is a $(T \times n)$ matrix with rows Y_t', $\mathbf{X}$ is a $(T \times k)$ matrix ($k = 1 + np, p$ =number of lags) with rows $X_t' = [1, Y_{t-1}', ..., Y_{t-p}']$, $\mathbf{U}$ is a $(T \times n)$ matrix with rows u_t' and Φ is a $(k \times n) = [\Phi_0, \Phi_1, ..., \Phi_p]'$.The one-step-ahead forecast errors u_t have a multivariate normal distribution $N(0, \Sigma_u)$ conditional on past observations of Y. The log-likelihood function of the data is a function of Φ and Σ_u

$$L(\mathbf{Y}|\Phi, \Sigma_u) \propto |\Sigma_u|^{-\frac{T}{2}} \exp\left\{ -\frac{1}{2} tr\left[\Sigma_u^{-1} \left(\mathbf{Y'Y} - \mathbf{\Phi'X'Y} - \mathbf{Y'X\Phi} + \mathbf{\Phi'X'X\Phi} \right) \right] \right\} \tag{18}$$

The prior distribution for the VAR parameters proposed by Del Negro and Schorfheide (2004) is based on the statistical representation of the DSGE model given by a VAR approximation. Let Γ_{xx}^*, Γ_{yy}^*, Γ_{xy}^* and Γ_{yx}^* be the theoretical second-order moments of the variables Y and X implied by the DSGE model, where

$$\begin{gathered} \Phi^*(\theta) = \Gamma_{xx}^{*-1}(\theta) \Gamma_{xy}^*(\theta) \\ \Sigma^*(\theta) = \Gamma_{yy}^*(\theta) - \Gamma_{yx}^*(\theta) \Gamma_{xx}^{*-1}(\theta) \Gamma_{xy}^*(\theta) \end{gathered} \tag{19}$$

The moments are the dummy observation priors used in the mixture model. These vectors can be interpreted as the probability limits of the coefficients in a VAR estimated on the artificial observations generated by the DSGE model. Conditional on the vector of structural parameters in the DSGE model θ, the prior distributions for the VAR parameters $p(\Phi, \Sigma_u|\theta)$ are of the Inverted-Wishart (IW) and Normal forms

$$\begin{gathered} \Sigma_u | \theta \sim IW\left((\lambda T \Sigma_u^*(\theta), \lambda T - k, n \right) \\ \mathbf{\Phi} | \Sigma_u, \theta \sim N\left(\mathbf{\Phi}^*(\theta), \Sigma_u \otimes (\lambda T \Gamma_{XX}(\theta))^{-1} \right) \end{gathered} \tag{20}$$

where the parameter λ controls the degree of model misspecification with respect to the VAR: for small values of λ the discrepancy between the VAR and the DSGE-VAR is large and a sizeable distance is generated between the unrestricted VAR and DSGE estimators. Large values of λ correspond to small model misspecification and for $\lambda = \infty$ beliefs about DSGE misspecification degenerate to a point mass at zero. Bayesian estimation could be interpreted as estimation based on a sample in which data are augmented by a hypothetical sample in which observations are generated by the DSGE model, the so-called dummy prior observations (Theil and Goldberg, 1961; Ingram and Whiteman, 1994). Within this framework λ determines the length of the hypothetical sample. The posterior distributions of the VAR parameters are also of the Inverted-Wishart and Normal forms. Given the prior distribution, posterior distributions are derived by the Bayes theorem

$$\Sigma_u \left|\theta, \mathbf{Y} \sim IW\left(\left(\lambda+1\right) T \hat{\Sigma}_{u,b}\left(\theta\right), \left(\lambda+1\right) T - k, n \right)\right. \tag{21}$$

$$\mathbf{\Phi} \left|\Sigma_u, \theta, \mathbf{Y} \sim N\left(\hat{\mathbf{\Phi}}_b\left(\theta\right), \Sigma_u \otimes \left[\lambda T \Gamma_{XX}\left(\theta\right) + \mathbf{X}'\mathbf{X}\right]^{-1} \right)\right. \tag{22}$$

$$\hat{\mathbf{\Phi}}_b\left(\theta\right) = \left(\lambda T \Gamma_{XX}\left(\theta\right) + \mathbf{X}'\mathbf{X}\right)^{-1} \left(\lambda T \Gamma_{XY}\left(\theta\right) + \mathbf{X}'\mathbf{Y}\right) \tag{23}$$

$$\hat{\Sigma}_{u,b}\left(\theta\right) = \frac{1}{\left(\lambda+1\right)T} \left[\left(\lambda T \Gamma_{YY}\left(\theta\right) + \mathbf{Y}'\mathbf{Y}\right) - \left(\lambda T \Gamma_{XY}\left(\theta\right) + \mathbf{X}'\mathbf{Y}\right) \hat{\mathbf{\Phi}}_b\left(\theta\right) \right] \tag{24}$$

where the matrices $\hat{\mathbf{\Phi}}_b\left(\theta\right)$ and $\hat{\Sigma}_{u,b}\left(\theta\right)$ have the interpretation of maximum likelihood estimates of the VAR parameters based on the combined sample of actual observations and artificial observations generated by the DSGE. Equations (21) and (22) show that the smaller λ is, the closer the estimates are to the OLS estimates of an unrestricted VAR. Instead, the higher λ is, the closer the VAR estimates will be tilted towards the parameters in the VAR approximation of the DSGE model ($\hat{\mathbf{\Phi}}_b\left(\theta\right)$ and $\hat{\Sigma}_{u,b}\left(\theta\right)$). In order to obtain a non-degenerate prior density (20), which is a necessary condition for the existence of a well-defined Inverse-Wishart distribution and for computing meaningful marginal likelihoods, λ has to be greater than λ_{MIN}

$$\begin{aligned} \lambda_{MIN} &\geq \frac{n+k}{T}; k = 1 + p \times n \\ p &= \text{lags} \\ n &= \text{endogenous variables.} \end{aligned}$$

Hence, the optimal lambda must be greater than or equal to the minimum lambda $\left(\hat{\lambda} \geq \lambda_{MIN}\right)$. Essentially, the DSGE-VAR tool allows the econometrician to draw posterior inferences about the DSGE model parameters θ. Del Negro and Schorfheide (2004) explain that the posterior estimate of θ has the interpretation of a minimum-distance estimator, where the discrepancy between the OLS estimates of the unrestricted VAR parameters and the VAR representation of the DSGE model is a sort of distance function. The estimated posterior of parameter vector θ depends on the hyperparameter λ. When $\lambda \to 0$, in the posterior the parameters are not informative, so the DSGE model is of no use in explaining the data. Unfortunately, the posteriors (22) and (21) do not have a closed form and we need a numerical method to solve the problem. The posterior simulator used by Del Negro and Schorfheide (2004) is the Markov Chain Monte Carlo Method and the algorithm used is the Metropolis-Hastings acceptance method. This procedure generates a Markov Chain from the posterior distribution of θ and this Markov Chain is used for Monte Carlo simulations. The optimal λ is given by maximizing the log of the marginal data density

$$\hat{\lambda} = \underset{\lambda \geqslant \lambda_{MIN}}{\arg\max} \ln p(\mathbf{Y}|\lambda)$$

According to the optimal lambda $\left(\hat{\lambda}\right)$, a corresponding optimal mixture model is chosen. This hybrid model is called DSGE-VAR$\left(\hat{\lambda}\right)$ and $\hat{\lambda}$ is the weight of the priors. It can also be interpreted as the restriction of the theoretical model on the actual data.

3.3 Factor Augmented DSGE (DSGE-FAVAR)

Following Bernanke *et al.* (2005), a FAVAR benchmark for the evaluation of a DSGE model will include a vector of observable variables and a small vector of unobserved factors extracted from a large data-set of macroeconomic time series, that capture additional economic information relevant to model the dynamics of the observables. In this study we implement the DSGE-FAVAR model of Consolo *et al.* (2009). The statistical representation has the following specification:

$$\begin{aligned} \begin{pmatrix} \mathbf{Y}_t \\ \mathbf{F}_t \end{pmatrix} &= \begin{bmatrix} \boldsymbol{\Phi}_{11}(L) & \boldsymbol{\Phi}_{12}(L) \\ \boldsymbol{\Phi}_{21}(L) & \boldsymbol{\Phi}_{22}(L) \end{bmatrix} \begin{pmatrix} \mathbf{Y}_{t-1} \\ \mathbf{F}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{u}_t^Y \\ \mathbf{u}_t^F \end{pmatrix} \\ \mathbf{Y}_t &= (\Delta \ln x_t, \Delta \ln P_t, \ln R_t, SP_t) \\ \mathbf{F}_t &= \left(F_{1t}^s, F_{2t}^s, F_{3t}^f\right) \end{aligned} \tag{25}$$

where $\mathbf{Y}_t$ are the observable variables included in the simple DSGE model and $\mathbf{F}_t$ is a small vector of unobserved factors relevant to modelling the dynamics of $\mathbf{Y}_t$ (F_{1t}^s, F_{2t}^s are the two slow factors and F_{3t}^f is the fast factor). The system reduces to the standard VAR when $\Phi_{12}(L) = 0$. Importantly, and differently from Boivin and Giannoni (2006), this FAVAR is not interpreted as the reduced form of a DSGE model at hand. In fact, in this case the restrictions implied by the DSGE model on a general FAVAR are very difficult to trace and model evaluation becomes even more difficult to implement. A very tightly parameterized theory model can have a very highly parameterized reduced form if one is prepared to accept that the relevant theoretical concepts in the model are a combination of many macroeconomic and financial variables. The DSGE-FAVAR is implemented in the same way as the DSGE-VAR.

4 Multivariate State-Space Time-Varying Parameter VAR models

Time varying parameter autoregression could easily form a state space model with the parameters of the TVP-VAR as state variables. The state space model has been well studied by Harvey (1990) and Durbin and Koopman (2002). According to Kalman (1960, 1963), in a state-space representation the signal extraction is implemented through a model that links the unobserved and observed variables of the system. To estimate a state space model, several methods have been developed. Kalman filtering involves sequentially updating a linear projection on the vector of interest. The state-space representation is given by a system of two vector equations. Firstly, the state or transition equation describes the dynamics of the state vector containing the unobserved variables for estimation, while the second equation represents the observation or measurement equation linking the state vector to the vector containing the observed variables. For the standard homoskedastic TVP-VAR models, the parameters are estimated using a multivariate specification of the standard Kalman filter (Harvey, 1990; Bekiros and Paccagnini, 2013). The likelihood estimation requires repeating the filtering many times in order to evaluate the likelihood for each set of the time-varying parameters until the maximum is reached. This is performed with a suitable multivariate extension of the Kim and Nelson (1999a, 1999b) method. The calculation of the Hessian for the estimation of the variance-covariance matrix is done with the Broyden-Fletcher-Goldfarb-Shano (BFGS) optimization algorithm. Other algorithms can also be used with the same results, e.g., the DFP and the Levenberg-Marquardt. The parameters could be also estimated with the use of the Zellner g-prior both for homoskedastic and heteroskedastic TVP-VARs and in this case the numerical evaluation of the posterior distributions is performed with Gibbs sampling (Kim and Nelson, 1999b).

4.1 MVSS-TVP-VAR model with Homoscedastic errors (standard model)

The standard homoskedastic TVP-VAR (MVSS-TVP-VAR) can be expressed as

$$\mathbf{y}_t = \mathbf{B}_{0,t} + \mathbf{B}_{1,t}\mathbf{y}_{t-1} + \cdots + \mathbf{B}_{p,t}\mathbf{y}_{t-p} + u_t \tag{26}$$

in which $\mathbf{B}_{0,t}$ is a $k \times 1$ vector of time-varying intercepts, $\mathbf{B}_{i,t}$ $(i = 1, \ldots, p)$ are $k \times k$ matrices of time-varying coefficients and u_t are homoscedastic reduced-form residuals with a covariance matrix $\boldsymbol{\Omega}_t$. This could be transformed into a multivariate state-space form. First, consider the following state-space system:

$$y_t = \mathbf{Z}_t \boldsymbol{\alpha}_t + \varepsilon_t \tag{27}$$

$$\boldsymbol{\alpha}_t = \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t \tag{28}$$

The first equation is known as the *measurement* or *observation* equation and presents that part of the system than can physically be measured, while the second is the *state* equation $\boldsymbol{\alpha}_t$ the vector of state variables. The variables in this equation may or may not be observable. In the case at hand, they are not observable but will be estimated by the Kalman filter. $\mathbf{Z}_t$ is a matrix of known or unknown time varying coefficients and matrix $\mathbf{T}_t$, the state transition matrix. Finally, ε_t is $N\left(0,\sigma^2\right)$ while $\boldsymbol{\eta}_t$ in multivariate normal with an expected value of zero and a homoscedastic covariance matrix of $\mathbf{Q}$. The unknown parameters - called *hyperparameters* - are the elements of the matrices and the variances of the noise processes to be estimated. This is accomplished by maximizing the likelihood function which is presented below for one time period

$$L_t = -\frac{1}{2}\sum_{t=1}^{T}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}\ln f_t - \frac{1}{2}\sum_{t=1}^{T}\frac{\eta_t^2}{f_t} \tag{29}$$

where η_t is the one–step ahead residual at time t and f_t is its variance. It is calculated recursively using the following equations:

$$\boldsymbol{\alpha}_{t|t-1} = \mathbf{T}_t\boldsymbol{\alpha}_t \tag{30}$$

$$\mathbf{P}_{t|t-1} = \mathbf{T}_t\mathbf{P}_t\mathbf{T}_t' + \mathbf{Q} \tag{31}$$

$$\eta_t = y_t - \mathbf{Z}_t\boldsymbol{\alpha}_{t|t-1} \tag{32}$$

$$f_t = \mathbf{Z}_t\mathbf{P}_{t|t-1}\mathbf{Z}_t' + \sigma^2 \tag{33}$$

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{Z}_t'\eta_t/f_t \tag{34}$$

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{Z}_t'\mathbf{Z}_t\mathbf{P}_{t|t-1}/f_t \tag{35}$$

Hence, Equations (30) to (35) that generate an estimate of the state vector and its covariance matrix $\mathbf{P}_t$ are known as the Kalman filter. Given starting values, an estimate of the unknown regression coefficients is obtained. Then using this information in the likelihood function, one may then estimate the hyperparameters of the model. Once these estimates have been obtained, an estimate of the state vector, the recursive residuals and their variance is obtained, and also an estimate of the updated residual vector $e_t = y_t - \mathbf{Z}_t\boldsymbol{\alpha}_t$ is generated.

Harvey (1990) provides a framework for a multivariate version of the Kalman filter based on a time series analogue of the seemingly unrelated regression equation (SURE) model introduced into econometrics by Zellner (1963). Harvey (1990) refers to it as a system of seemingly unrelated time series equations (SUTSE) model. The simplest SUTSE model is the multivariate random walk plus noise process:

$$\mathbf{y}_t = \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \quad t = 1,\ldots,T \tag{36}$$

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t \tag{37}$$

where $\boldsymbol{\alpha}_t$ is an $N\times 1$ vector of local level components and $\boldsymbol{\varepsilon}_t$, and $\boldsymbol{\eta}_t$ are vectors of multivariate white noise with mean zero and covariance matrices $\boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{\Sigma}_\eta$ respectively. As in the univariate model, $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ are assumed to be uncorrelated with each other in all time periods. The variables are linked via the off-diagonal elements $\boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{\Sigma}_\eta$. An important property of the SUTSE system is that its form remains unaltered when it is subject to contemporaneous aggregation. A linear time-invariant univariate structural model can be written in the SUTSE state space form for N variables

$$\mathbf{y}_t = \left(\mathbf{z}'\otimes\mathbf{I}_N\right)\boldsymbol{\alpha}_t + \varepsilon_t \tag{38}$$

$$\boldsymbol{\alpha}_t = \left(\mathbf{T}\otimes\mathbf{I}_N\right)\boldsymbol{\alpha}_{t-1} + \left(\mathbf{R}\otimes\mathbf{I}_N\right)\boldsymbol{\eta}_t \tag{39}$$

with $Var(\boldsymbol{\varepsilon}_t) = \boldsymbol{\Sigma}_\varepsilon$ and $Var(\boldsymbol{\eta}_t)$ a block diagonal matrix with the blocks being $\boldsymbol{\Sigma}_k$, $k = 1,...,g$. For example, in the four-variate case (as in this study) the variance of the error component in the state equation is

$$Var\left(\boldsymbol{\eta}_t\right) = \begin{bmatrix} \boldsymbol{\Sigma}_\eta & 0 & 0 & 0 \\ 0 & \boldsymbol{\Sigma}_\zeta & 0 & 0 \\ 0 & 0 & \boldsymbol{\Sigma}_\omega & 0 \\ 0 & 0 & 0 & \boldsymbol{\Sigma}_\gamma \end{bmatrix} \tag{40}$$

In fact a more general formulation of the SUTSE model does not constrain $Var(\boldsymbol{\eta}_t)$ to be diagonal and hence $Var(\boldsymbol{\eta}_t)$ need not be block diagonal. Indeed the SUTSE formulation can be generalized further to allow quantities such as $\mathbf{z}$,$\boldsymbol{\Sigma}_\varepsilon$, $\mathbf{T}$, $\mathbf{R}$ and $Var(\boldsymbol{\eta}_t)$ to change deterministically over time. As shown in Harvey (1986), the time-domain treatment still goes through. The Kalman filter may be applied to (38) and (39), the number of sets of observations needed to form an estimator of $\boldsymbol{\alpha}_t$, with finite MSE matrix being the same as in the univariate case. The conditions for the filter to converge to a steady state are an obvious generalization of the conditions in the univariate case. Given normality of the disturbances, the log-likelihood function is of the prediction error decomposition form.

The decoupling of the Kalman filter is related to the result which arises in a SURE system where OLS applied to each equation in turn is fully efficient if each equation contains the same regressors. Hence, all the information needed for estimation, prediction and smoothing can be obtained by applying the same univariate filter to each series in turn. Consider the multivariate random walk plus noise model. If the signal-to-noise ratio is q (i.e., $\boldsymbol{\Sigma}_\eta/\boldsymbol{\Sigma}_\varepsilon = q$), the Kalman filter for this model is

$$\boldsymbol{\alpha}_{t+1|t} = \boldsymbol{\alpha}_{t|t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \boldsymbol{\alpha}_{t|t-1}\right), t = 2, \ldots, T \tag{41}$$

and

$$\mathbf{P}_{t+1|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{F}_t^{-1}\mathbf{P}_{t|t-1} + q\boldsymbol{\Sigma}_\varepsilon \tag{42}$$

where

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{F}_t^{-1} \tag{43}$$

and

$$\mathbf{F}_t = \mathbf{P}_{t|t-1} + \boldsymbol{\Sigma}_\varepsilon \tag{44}$$

Let w_t denote a positive scalar for $t = 2, \ldots, T$ and suppose that $\mathbf{P}_{t|t-1}$, the MSE matrix of the $N \times 1$ vector $\boldsymbol{\alpha}_{t|t-1}$, is proportional to $\boldsymbol{\Sigma}_\varepsilon$, i.e. $\mathbf{P}_{t|t-1} = w_t\boldsymbol{\Sigma}_\varepsilon$. It then follows from (42) that $\mathbf{P}_{t+1|t}$ is of the same form, that is, $\mathbf{P}_{t+1|t} = w_{t+1}\boldsymbol{\Sigma}_\varepsilon$ with $w_{t+1} = (w_t + w_t q + q)/(w_t + 1)$. Furthermore if $\mathbf{P}_{t|t-1} = w_t\boldsymbol{\Sigma}_\varepsilon$ the gain matrix in (41) is diagonal, that is

$$\mathbf{K}_t = w_t\boldsymbol{\Sigma}_\varepsilon\left(w_t\boldsymbol{\Sigma}_\varepsilon + \boldsymbol{\Sigma}_\varepsilon\right)^{-1} = \left[w_t/\left(w_{t+1}\right)\right]\mathbf{I}_N \tag{45}$$

Suppose that the above Kalman filter is started off in such a way that $\mathbf{P}_{2|1}$ is proportional to $\boldsymbol{\Sigma}_\varepsilon$; that is $\mathbf{P}_{2|1} = p_{_{2|1}}\boldsymbol{\Sigma}_\varepsilon$, where $p_{_{2|1}}$ is a scalar. Since $\mathbf{P}_{t|t-1}$ must continue to be proportional to $\boldsymbol{\Sigma}_\varepsilon$, it follows from (45) that the elements of $\boldsymbol{\alpha}_{t+1|t}$, can be computed from the univariate recursions. It also follows that w_t, must be equal to $p_{_{t|t-1}}$ for all $t = 2, \ldots, T$. The starting values $\boldsymbol{\alpha}_{_{2|1}} = \mathbf{y}_1$ and $\mathbf{P}_{2|1} = \boldsymbol{\Sigma}_\eta + \boldsymbol{\Sigma}_\varepsilon = (1+q)\boldsymbol{\Sigma}_\varepsilon$ equally correspond to the use of a diffuse prior, and the use of these starting values leads to the exact likelihood function for $\mathbf{y}_2, \ldots, \mathbf{y}_T$ in the prediction error decomposition form

$$\log L = -\frac{(T-1)N}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{T}\log|\mathbf{F}_t| - \frac{1}{2}\sum_{t=2}^{T}\mathbf{v}_t'\mathbf{F}_t^{-1}\mathbf{v}_t \tag{46}$$

where $\mathbf{v}_t = \mathbf{y}_t - \tilde{\mathbf{y}}_{t|t-1}, t = 1, ..., T$. However, the decoupling of the Kalman filter allows the elements of $\mathbf{v}_t$, to be computed from the univariate recursions. Furthermore

$$\mathbf{P}_{t|t-1} = p_{_{t|t-1}}\boldsymbol{\Sigma}_\varepsilon \tag{47}$$

and so

$$\mathbf{F}_t = \mathbf{P}_{t|t-1} + \boldsymbol{\Sigma}_\varepsilon = f_t\boldsymbol{\Sigma}_\varepsilon,\ t = 3, \ldots, T \tag{48}$$

where $f_t = \left(p_{_{t|t-1}} + 1\right)$. Substituting from (48) into (46) gives

$$\log L = -\frac{(T-1)N}{2}\log 2\pi + \frac{(T-1)}{2}\log\left|\boldsymbol{\Sigma}_\varepsilon^{-1}\right| - \frac{N}{2}\sum_{t=2}^{T}\log f_t - \frac{1}{2}\sum_{t=2}^{T}\frac{1}{f_t}\mathbf{v}_t'\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{v}_t \tag{49}$$

Differentiating (49) with respect to the distinct elements of $\boldsymbol{\Sigma}_\varepsilon^{-1}$ leads to the ML estimator of $\boldsymbol{\Sigma}_\varepsilon$ being

$$\tilde{\mathbf{\Sigma}}_{\varepsilon} = (T-1)^{-1} \sum_{t=2}^{T} f_t^{-1} \mathbf{v}_t \mathbf{v}_t' \tag{50}$$

for any given value of q. The ML estimators of q and $\mathbf{\Sigma}_{\varepsilon}$ can therefore be obtained by maximizing the concentrated likelihood function

$$\log L_c = -\frac{(T-1)\,N}{2} \log 2\pi - \frac{(T-1)}{2} \log \left| \tilde{\mathbf{\Sigma}}_{\varepsilon} \right| - \frac{N}{2} \sum_{t=2}^{T} \log f_t \tag{51}$$

with respect to q. Once the parameters have been estimated, prediction and smoothing can be carried out. The predictions of future observations are obtained from the univariate recursions

$$\text{MSE}\left(\tilde{\mathbf{y}}_{T+l|T}\right) = f_{T+l|T} \mathbf{\Sigma}_{\varepsilon},\ l = 1, 2, \ldots \tag{52}$$

where

$$f_{T+l|T} = p_{T+l|T} + 1 \tag{53}$$

The decoupling of the Kalman filter can be shown in a similar way for the time-varying system as in Bekiros and Paccagnini (2013):

$$\mathbf{y}_t = (\mathbf{z}_t' \otimes \mathbf{I}_N)\,\boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, Var\left(\boldsymbol{\varepsilon}_t\right) = h_t \mathbf{\Sigma}_* \tag{54}$$

$$\boldsymbol{\alpha}_t = (\mathbf{T}_t \otimes \mathbf{I}_N)\,\boldsymbol{\alpha}_{t-1} + (\mathbf{R}_t \otimes \mathbf{I}_N)\,\boldsymbol{\eta}_t,\ Var\left(\boldsymbol{\eta}_t\right) = \mathbf{Q}_t \otimes \mathbf{\Sigma}_* \tag{55}$$

where $\mathbf{Q}_t$ =diag$(q_1, \ldots, q_k)$. The more general formulation does not constrain $\mathbf{Q}_t$ to be diagonal, although, as in the univariate model, restrictions are needed on $\mathbf{Q}_t$ for the model to be identifiable. All the results on estimation and prediction carry through, with $\mathbf{P}_{t+1|t} = \mathbf{P}^*_{t+1|t} \otimes \mathbf{\Sigma}_*$, where $\mathbf{P}^*_{t+1|t}$ is the MSE matrix for the univariate model (Harvey, 1986, 1990).

4.2 MVSS-TVP-VAR model with Markov-Switching Heteroscedasticity

As an alternative to the homoskedastic TVP-VAR we assume that ε_t and $\boldsymbol{\eta}_t$ (i.e., $\sigma^2$and $\mathbf{Q}$) are dependent on Hamilton's (1988) state variable (S_t), which is an outcome of an unobserved discrete-time, discrete-state Markov process. While the Markov-switching variance model fails to incorporate the learning process of agents, the TPV model fails to incorporate uncertainty that changes due to future random shocks. In this section we consider a more general model in which both types of uncertainty are incorporated. An important motivation for considering a state-space model with Markov-switching heteroskedasticity is due to Lastrapes (1989), Lamoureux and Lastrapes (1990) and Kim (1993) who showed that failure to allow for regime shifts leads to an overstatement of the persistence of the variance of a series. Moreover, in this way we could incorporate different regimes in crisis periods.

Consider the following first-order, ω-state Markov-switching model of heteroskedasticity:

$$Q_t = Q^{S_t} = Q_1 \Theta_{1t} + Q_2 \Theta_{2t} + \cdots + Q_\omega \Theta_{\omega t} \tag{56}$$

$$h_t = h^{S_t} = h_1 \Theta_{1t} + h_2 \Theta_{2t} + \cdots + h_\omega \Theta_{\omega t} \tag{57}$$

where $\Theta_{jt} = 1$ if $S_t = j$ and $\Theta_{jt} = 0$ if $S_t \neq j\ (j = 1, 2, \ldots, \omega)$. The unobserved-state variable S_t evolves according to a Markov process with transition probability

$$p = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1\omega} \\ p_{21} & p_{22} & \cdots & p_{2\omega} \\ \vdots & \vdots & \ddots & \vdots \\ p_{\omega 1} & p_{\omega 2} & \cdots & p_{\omega\omega} \end{pmatrix} \tag{58}$$

where $p_{ij} = \Pr\left[S_t = j | S_{t-1} = i\right]$, for $i, j = 1, 2, \ldots, \omega$, and $\sum\limits_{j=1}^{\omega} p_{ij} = 1$.

Under a state-space form we consider the time-varying-parameter model in Eqs. (27)-(28), in this case with a 2-state Markov-switching model of heteroskedasticity:

$$\begin{aligned}\Pr\left[S_t = 1|S_{t-1}=1\right] &= P_{11};\ \Pr\left[S_t = 0|S_{t-1}=1\right] = 1-P_{11};\\ \Pr\left[S_t = 1|S_{t-1}=0\right] &= 1-P_{00};\ \Pr\left[S_t = 0|S_{t-1}=0\right] = P_{00}\end{aligned}$$

4.2.1 A Quasi-Optimal Filter for Approximation

Suppose that the parameters Q^j, h^j $(j=1,2)$, $\mathbf{Z}_t$, $\mathbf{T}_t$, and p are known for the model with a 2-state Markov-switching heteroskedasticity that consists of Equations (27)-(28) and (56)-(58). Given that $S_{t-1}=i$ and $S_t = j$ $(i,j=1,2)$, the Kalman filter can be represented as follows:

$$\boldsymbol{\alpha}^i_{t|t-1} = \mathbf{T}_t \boldsymbol{\alpha}^i_{t-1|t-1} \tag{59}$$

$$\mathbf{P}^{(i,j)}_{t|t-1} = \mathbf{T}_t \mathbf{P}^i_{t-1|t-1}\mathbf{T}'_t + Q^j \tag{60}$$

$$\eta^i_{t|t-1} = y_t - \mathbf{Z}_t\boldsymbol{\alpha}^i_{t|t-1} \tag{61}$$

$$f^{(i,j)}_{t|t-1} = \mathbf{Z}_t\mathbf{P}^{(i,j)}_{t|t-1}\mathbf{Z}'_t + h^j \tag{62}$$

$$\boldsymbol{\alpha}^{(i,j)}_{t|t} = \mathbf{T}_t\boldsymbol{\alpha}^i_{t|t-1} + K^{(i,j)}_t\eta^j_{t|t-1} \tag{63}$$

$$\mathbf{P}^{(i,j)}_{t|t} = \left(I - K^{(i,j)}_t\mathbf{Z}_t\right)\mathbf{P}^{(i,j)}_{t|t-1} \tag{64}$$

and

$$K^{(i,j)}_t = \mathbf{P}^{(i,j)}_{t|t-1}\mathbf{Z}_t{}'\left(f^{(i,j)}_{t|t-1}\right)^{-1} \tag{65}$$

where $\boldsymbol{\alpha}^i_{t|t-1}$ is an inference of $\boldsymbol{\alpha}_t$ based on information up to time $t-1$, given $S_{t-1}=i$; $\mathbf{P}^{(i,j)}_{t|t-1}$ is the covariance matrix of $\boldsymbol{\alpha}^i_{t|t-1}$; $\eta^i_{t|t-1}$ is the conditional forecast error of y_t based on information up to time $t-1$, given $S_{t-1}=i$; $f^{(i,j)}_{t|t-1}$ is the conditional variance of forecast error $\eta^i_{t|t-1}$, given $S_{t-1}=i$ and $S_t=j$; and $K^{(i,j)}_t$ is the Kalman gain, given $S_{t-1}=i$ and $S_t=j$. If we iterated the preceding algorithm from $t=1$ to $t=T$, the inferences on $\boldsymbol{\alpha}_T$ and its covariance matrix $\left(\boldsymbol{\alpha}_{T|T}\text{ and }\mathbf{P}_{T|T}\right)$ for example, would depend on the whole history of current and past states, $S_0, S_1, \ldots, S_T$. In general, we would have ω^T cases to consider, which would be quite impossible to deal even with relatively few observations. In other words, each iteration of the preceding filter produces an ω-fold increase in the number of cases to consider, as in the work of Harrison and Stevens (1976) and Gordon and Smith (1988). In this case, we have $\omega^T = 2^T$ cases to consider. However, we would like to reduce the dimension of the posteriors in (63) and (64) into (2×1) at the end of each iteration, hence we would have only (2×2) cases to consider for each iteration. Consider the following approximations employed for this purpose; If $\boldsymbol{\alpha}^{(i,j)}_{t|t}$ in (63) represented $E\left[\boldsymbol{\alpha}_t|S_{t-1}=i, S_t=j, \psi_t\right]$, it would be straightforward to show that

$$\boldsymbol{\alpha}^j_{t|t} = \sum_{i=1}^{2}\zeta_t\boldsymbol{\alpha}^{(i,j)}_{t|t} \tag{66}$$

where $\zeta_t = \Pr\left[S_{t-1}=i, S_t=j|\psi_t\right]\ /\ \Pr\left[S_t=j|\psi_t\right]$ and $\boldsymbol{\alpha}^j_{t|t}$ would represent $E\left[\boldsymbol{\alpha}_t|S_t=j,\psi_t\right]$. In this case, the covariance matrix of $\boldsymbol{\alpha}_t$ conditional on ψ_t and on $S_t=j$ could be derived in the following way:

$$
\begin{aligned}
\mathbf{P}_{t|t}^{j} &= E\left[(\boldsymbol{\alpha}_t - E\left[\boldsymbol{\alpha}_t|S_t=j,\psi_t\right]) \times (\boldsymbol{\alpha}_t - E\left[\boldsymbol{\alpha}_t|S_t=j,\psi_t\right])' \,|S_t=j,\psi_t\right] = E\left[\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{j}\right)\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{j}\right)'|S_t=j,\psi_t\right] \\
&= \sum_{i=1}^{2}\zeta_t E\left[\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{j}\right)\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{j}\right)'|S_{t-1}=i,S_t=j,\psi_t\right] \\
&= \sum_{i=1}^{2}\zeta_t E\left[\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}+\boldsymbol{\alpha}_{t|t}^{(i,j)}-\boldsymbol{\alpha}_{t|t}^{j}\right)\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}+\boldsymbol{\alpha}_{t|t}^{(i,j)}-\boldsymbol{\alpha}_{t|t}^{j}\right)'|S_{t-1}=i,S_t=j,\psi_t\right] \\
&= \sum_{i=1}^{2}\zeta_t \left\{E\left[\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)'|S_{t-1}=i,S_t=j,\psi_t\right]+\left(\boldsymbol{\alpha}_{t|t}^{j}-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)\left(\boldsymbol{\alpha}_{t|t}^{j}-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)'\right\} \\
&\quad +\sum_{i=1}^{2}\zeta_t\left(E\left[\boldsymbol{\alpha}_t|S_{t-1}=i,S_t=j,\psi_t\right]-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)\left(\boldsymbol{\alpha}_{t|t}^{(i,j)}-\boldsymbol{\alpha}_{t|t}^{j}\right)'+ \\
&\quad +\sum_{i=1}^{2}\zeta_t\left(\boldsymbol{\alpha}_{t|t}^{(i,j)}-\boldsymbol{\alpha}_{t|t}^{j}\right)\left(E\left[\boldsymbol{\alpha}_t|S_{t-1}=i,S_t=j,\psi_t\right]-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)' \\
&\quad =\sum_{i=1}^{2}\zeta_t \left\{E\left[\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)'|S_{t-1}=i,S_t=j,\psi_t\right]+\left(\boldsymbol{\alpha}_{t|t}^{j}-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)\left(\boldsymbol{\alpha}_{t|t}^{j}-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)'\right\} \qquad (67)
\end{aligned}
$$

Here, if $\mathbf{P}_{t|t}^{(i,j)}$ in (64) represented $E\left[\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)'|S_{t-1}=i,S_t=j,\psi_t\right]$ then (67) could be rewritten as

$$
\mathbf{P}_{t|t}^{j}=\sum_{i=1}^{\omega}\zeta_t\left\{\mathbf{P}_{t|t}^{(i,j)}+\left(\boldsymbol{\alpha}_{t|t}^{j}-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)\left(\boldsymbol{\alpha}_{t|t}^{j}-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)'\right\} \qquad (68)
$$

At the end of each iteration, we employ Equations (66) and (68) to "collapse" 2×2 posteriors in (63) and (64) into 2×1 to make the filter appropriate. Notice, however that these collapsed posteriors involve approximation because $\boldsymbol{\alpha}_{t|t}^{(i,j)}$ and $\mathbf{P}_{t|t}^{(i,j)}$ in (63) and (64) do not calculate $E\left[\boldsymbol{\alpha}_t|S_{t-1}=i,S_t=j,\psi_t\right]$ and $E\left[\begin{array}{c}\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)\times\left(\boldsymbol{\alpha}_t-\boldsymbol{\alpha}_{t|t}^{(i,j)}\right)'| \\ S_{t-1}=i,S_t=j,\psi_t\end{array}\right]$ exactly. This is because $\boldsymbol{\alpha}_t$ conditional on $\psi_{t-1}, S_t=j$ and $S_{t-1}=i$ is a mixture of Normals for $t>2$. It is because this approximation is employed that the preceding filter is called a *quasi-optimal filter*. The last thing that remains to be considered to complete the filter is to calculate $\Pr\left[S_{t-1}=i,S_t=j|\psi_t\right]$ and other probability terms. We follow Hamilton (1989) with a slight modification:

$$
\begin{aligned}
\Pr\left[S_{t-1}=i,S_t=j|\psi_t\right] &= \frac{\Pr\left[y_t,S_{t-1}=i,S_t=j|\psi_{t-1}\right]}{\Pr\left[y_t|\psi_{t-1}\right]} \\
&= \frac{\Pr\left[y_t|S_{t-1}=i,S_t=j|\psi_{t-1}\right]\times\Pr\left[S_{t-1}=i,S_t=j|\psi_{t-1}\right]}{\Pr\left[y_t|\psi_{t-1}\right]} \qquad (69)
\end{aligned}
$$

where

$$
\Pr\left[y_t|S_{t-1}=i,S_t=j|\psi_{t-1}\right]=\frac{1}{\sqrt{2\pi f_{t|t-1}^{(i,j)}}}e^{\left\{-\frac{\left(\eta_{t|t-1}^{i}\right)^2}{2f_{t|t-1}^{(i,j)}}\right\}} \qquad (70)
$$

$$
\Pr\left[y_t|\psi_{t-1}\right]=\sum_{i=1}^{2}\sum_{j=1}^{2}\Pr\left[y_t,S_{t-1}=i,S_t=j|\psi_{t-1}\right] \qquad (71)
$$

and

$$
\Pr\left[S_{t-1}=i,S_t=j|\psi_{t-1}\right],\;\; \Pr\left[S_t=j|S_{t-1}=i\right]\Pr\left[S_{t-1}=i|\psi_{t-1}\right] \qquad (72)
$$

with

$$\Pr\left[S_{t-1}=i|\psi_{t-1}\right]=\sum_{S_{t-2}=1}^{2}\Pr\left[S_{t-2}=s_{t-2},S_{t-1}=i|\psi_{t-1}\right] \tag{73}$$

Thus Equations (59)-(73) complete the *quasi-optimal filter*. As a by-product of the quasi-optimal filter, the *approximated conditional log-likelihood function* can be obtained from (71):

$$\log L=\log\left(\Pr\left[y_T,y_{t-1},\ldots|\psi_0\right]\right)=\sum_{t=1}^{T}\log\left(\Pr\left[y_t|\psi_{t-1}\right]\right) \tag{74}$$

The preceding filter is derived under the assumption that the parameters of the underlying model are known. To estimate the parameters of the model, we can maximize the log-likelihood function in Equation (74) with respect to the underlying unknown parameters of the model.

Similarly to the homoskedastic TVP-VAR, the above model can be written in the SUTSE (multivariate) state space form for N variables

$$\mathbf{y}_t=\left(\mathbf{z}'\otimes\mathbf{I}_N\right)\boldsymbol{\alpha}_t+\varepsilon_t \tag{75}$$

$$\boldsymbol{\alpha}_t=\left(\mathbf{T}\otimes\mathbf{I}_N\right)\boldsymbol{\alpha}_{t-1}+\left(\mathbf{R}\otimes\mathbf{I}_N\right)\boldsymbol{\eta}_t \tag{76}$$

yet now $Var(\varepsilon_t)=Var(h_t^j)=\boldsymbol{\Sigma}_{h^j}$ and $Var(\boldsymbol{\eta}_t)=Var(Q_t^j)=\boldsymbol{\Sigma}_{Q^j}$ are block diagonal matrices with the blocks all of them following a 2-state Markov-switching heteroskedastic structure, namely in the four-variate case, the variance of the error components in the state equation is

$$Var\left(h_t^j\right)=\begin{bmatrix}\boldsymbol{\Sigma}_{1h^j} & 0 & 0 & 0\\ 0 & \boldsymbol{\Sigma}_{2h^j} & 0 & 0\\ 0 & 0 & \boldsymbol{\Sigma}_{3h^j} & 0\\ 0 & 0 & 0 & \boldsymbol{\Sigma}_{4h^j}\end{bmatrix} \tag{77}$$

and

$$Var\left(Q_t^j\right)=\begin{bmatrix}\boldsymbol{\Sigma}_{1Q^j} & 0 & 0 & 0\\ 0 & \boldsymbol{\Sigma}_{2Q^j} & 0 & 0\\ 0 & 0 & \boldsymbol{\Sigma}_{3Q^j} & 0\\ 0 & 0 & 0 & \boldsymbol{\Sigma}_{4Q^j}\end{bmatrix} \tag{78}$$

Indeed the SUTSE formulation can be generalized further to allow quantities such as $\mathbf{z},\mathbf{T},\mathbf{R}$ and $\boldsymbol{\Sigma}_{h^j},\boldsymbol{\Sigma}_{Q^j}$ to change deterministically over time. The quasi-Kalman filter for the heteroskedastic case may be applied to (75) and (76) with the number of sets of observations needed to form an estimator of $\boldsymbol{\alpha}_t$, and with finite MSE matrix being the same as in the univariate case. Moreover, the conditions for the filter to converge to a steady state define a generalization of the conditions in the univariate case. The decoupling of the Kalman filter is fully efficient if each equation contains the same regressors. Thus, all the information needed for estimation, prediction and smoothing can be obtained by applying the same univariate filter to each series in turn. If the signal-to-noise ratio is q (i.e., $\boldsymbol{\Sigma}_{Q^j}/\boldsymbol{\Sigma}_{h^j}=q$), the Kalman filter for this model is

$$\boldsymbol{\alpha}_{t+1|t}^{i}=\boldsymbol{\alpha}_{t|t-1}^{i}+\mathbf{K}_t^{(i,j)}\left(\mathbf{y}_t-\boldsymbol{\alpha}_{t|t-1}^{i}\right),t=2,\ldots,T \tag{79}$$

and

$$\mathbf{P}_{t+1|t}^{(i,j)}=\mathbf{P}_{t|t-1}^{(i,j)}-\mathbf{P}_{t|t-1}^{(i,j)}\left(\boldsymbol{\Sigma}_{Q^jt|t-1}^{(i,j)}\right)^{-1}\mathbf{P}_{t|t-1}^{(i,j)}+q\boldsymbol{\Sigma}_{h^j} \tag{80}$$

where

$$\mathbf{K}_t^{(i,j)}=\mathbf{P}_{t|t-1}^{(i,j)}\mathbf{Z}_t{}'\left(\boldsymbol{\Sigma}_{Q^jt|t-1}^{(i,j)}\right)^{-1} \tag{81}$$

and

$$\boldsymbol{\Sigma}_{Q^jt|t-1}^{(i,j)}=\mathbf{P}_{t|t-1}^{(i,j)}+\boldsymbol{\Sigma}_{h^j} \tag{82}$$

Let again w_t denote a positive scalar for $t = 2, \ldots, T$ and suppose that $\mathbf{P}^{(i,j)}_{t|t-1} = w_t \mathbf{\Sigma}_{h^j}$, i.e, the MSE matrix of the $N \times 1$ vector $\boldsymbol{\alpha}^i_{t|t-1}$, is proportional to $\mathbf{\Sigma}_{h^j}$ depending on the whole history of current and past states, $S_0, S_1, \ldots, S_T$. It follows from (80) that $\mathbf{P}^{(i,j)}_{t+1|t}$ is of the same form, that is, $\mathbf{P}^{(i,j)}_{t+1|t} = w_{t+1}\mathbf{\Sigma}_{h^j}$ with $w_{t+1} = (w_t + w_t q + q)/(w_t + 1)$, and if $\mathbf{P}^{(i,j)}_{t|t-1} = w_t \mathbf{\Sigma}_{h^j}$ the gain matrix in (79) is state-dependent diagonal, that is

$$\mathbf{K}^{(i,j)}_t = w_t \mathbf{\Sigma}_{h^j} \left(w_t \mathbf{\Sigma}_{h^j} + \mathbf{\Sigma}_{h^j}\right)^{-1} = \left[w_t/\left(w_{t+1}\right)\right] \mathbf{I}_N$$

In the heteroskedastic case the above Kalman filter is started off in the same way as in the standard model. However, the use of these starting values now would not lead to an exact likelihood function for $\mathbf{y}_2, \ldots, \mathbf{y}_T$ in the prediction error decomposition form as in (46), but now to a multivariate version of the approximated conditional log-likelihood function of the *quasi-optimal filter*

$$\mathbf{LL} = \log\left(\Pr\left[\mathbf{y}_T, \mathbf{y}_{t-1}, \ldots | \psi_0\right]\right) = \sum_{t=1}^{T} \log\left(\Pr\left[\mathbf{y}_t | \psi_{t-1}\right]\right) \tag{83}$$

based now on the following probabilities instead of the (70)-(71) for the univariate filter:

$$\Pr\left[\mathbf{y}_t | S_{t-1} = i, S_t = j | \psi_{t-1}\right] = \frac{1}{\sqrt{2\pi \mathbf{\Sigma}^{(i,j)}_{Q^j\, t|t-1}}} e^{\left\{-\frac{\left(f^i_{t|t-1}\right)^2}{2\mathbf{\Sigma}^{(i,j)}_{Q^j\, t|t-1}}\right\}} \tag{84}$$

$$\Pr\left[\mathbf{y}_t | \psi_{t-1}\right] = \sum_{i=1}^{2} \sum_{j=1}^{2} \Pr\left[\mathbf{y}_t, S_{t-1} = i, S_t = j | \psi_{t-1}\right] \tag{85}$$

The predictions of future observations are obtained from the univariate recursions (as in the standard model)

$$\text{MSE}\left(\tilde{\mathbf{y}}_{T+l|T}\right) = f_{T+l|T} \mathbf{\Sigma}_{h^j}, \; l = 1, 2, \ldots \text{and } p_{ij} = \Pr\left[S_t = j | S_{t-1} = i\right], i, j = 1, 2 \tag{86}$$

Finally, the decoupling of the Kalman filter can be extended in a similar way for a time-varying system with a 2-state Markov-switching model of heteroskedasticity, as in Bekiros and Paccagnini (2013):

$$\mathbf{y}_t = \left(\mathbf{z}'_t \otimes \mathbf{I}_N\right) \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \text{ with } \; Var(\varepsilon_t) = Var(h^j_t)\mathbf{\Sigma}_* = \mathbf{\Sigma}_{h^j} \mathbf{\Sigma}_* \tag{87}$$

$$\boldsymbol{\alpha}_t = \left(\mathbf{T}_t \otimes \mathbf{I}_N\right) \boldsymbol{\alpha}_{t-1} + \left(\mathbf{R}_t \otimes \mathbf{I}_N\right) \boldsymbol{\eta}_t, \text{ with } \; Var(\boldsymbol{\eta}_t) = Var(Q^j_t) \otimes \mathbf{\Sigma}_* = \mathbf{\Sigma}_{Q^j} \otimes \mathbf{\Sigma}_* \tag{88}$$

where again $Var(\varepsilon_t) = Var(h^j_t) = \mathbf{\Sigma}_{h^j}$ and $Var(\boldsymbol{\eta}_t) = Var(Q^j_t) = \mathbf{\Sigma}_{Q^j}$ are block diagonal matrices, although a more general formulation does not constrain them to be diagonal. However, as in the univariate model, restrictions are needed on the matrices for the model to be identifiable. The previous results on estimation and prediction apply, with $\mathbf{P}^{(i,j)*}_{t+1|t} = \mathbf{P}^{(i,j)}_{t+1|t} \otimes \mathbf{\Sigma}_*$, where $\mathbf{P}^{(i,j)}_{t+1|t}$ is the MSE matrix for the univariate model.

5 Empirical estimation

In this work we use quarterly data of the US economy from 1960:Q4 to 2010:Q4 with an out-of sample period that spans 1997:Q1 to 2010:Q4. The data for real output growth comes from the Bureau of Economic Analysis as Gross Domestic Product (GDP), while Consumer price index (CPI) data (seasonally adjusted, 1982-1984=100) are derived from the Bureau of Labor Statistics. Both series are taken in first difference logarithmic transformation. The interest rate series (FR rate) are constructed as in Clarida, Galì and Gertler (2000), namely for each quarter the interest rate is computed as the average federal funds rate during the first month of the quarter, including business days only. In the estimation of the DSGE model with financial frictions we follow Del Negro and Schorfheide (2012), namely we measure the Spread as the annualized Moody's Seasoned BAA Corporate Bond Yield spread over the 10-Year Treasury Note Yield at Constant Maturity[2].

The complete dataset is used to extract factors for FAVAR and DSGE-FAVAR models. In order to construct the FAVAR we extract factors from a balanced panel of 109 monthly and quarterly macroeconomic and financial time

[2] In this paper we do not consider a possible common stochastic trend as proposed by Del Negro and Schorfheide (2006) to avoid complications in the estimation procedure. We follow the standard DSGE-VAR model introduced in the seminal paper of Del Negro and Schorfheide (2004), where possible cointegration relations among the variables are neglected.

series, following the dataset built by Stock and Watson (2002). The dataset involves several measures of industrial production, interest rates, various price indices, employment and other important macroeconomic and also financial variables. In this set-up, the number of informational time series N is large (larger than time period T) and must be greater than the number of factors and observed variables in the FAVAR system ($k + M \ll N$). In the panel data used, there are some variables in monthly format, which are transformed into a quarterly data using end-of-period observations. All series have been transformed to induce stationarity. The series are taken as levels or transformed into logarithms, first or second difference (in level or logarithms) according to series characteristics [3]. Following Bernanke *et al.* (2005), we partition the data into two categories of information variables: slow and fast. Slow-moving variables (e.g., wages or spending) do not respond contemporaneously to unanticipated changes in monetary policy, while fast-moving variables (e.g., asset prices and interest rates) do respond contemporaneously to monetary shocks. Then we extract two factors from the slow variables and one factor from the fast variables. The methodology implemented to extract the factors is principal components. Stock-Watson (1998) showed that factors can be consistently estimated by the first r principal components of a matrix X, even in the presence of moderate changes in the loading matrix Λ. For this result to hold it is important that the estimated number of factors k, is larger than or equal to the true number, r. Bai and Ng (2000) propose a set of selection criteria to choose k that are generalizations of the BIC and AIC criteria. As they suggest, we use information criteria to determine the number of factors but, as they are not so decisive, we limit the number of factors to three to strike a balance between the variation in the original series explained by the principal components and the difference in the parameterization of the VAR and the FAVAR. It is also worth noting that the factors are not uniquely identified, but this is not a problem in our context because we will not attempt a structural interpretation of the estimated factors.

We compare the out-of-sample forecasting performance of VAR models including BVAR, FAVAR and the homoskedastic and heteroskedastic MVSS-TVP-VARs as well as of the DSGE class including DSGE-VAR, DSGE-FAVAR, in terms of the Root Mean Squared Forecast Error (RMSE) for the optimal lag specifications (one to four) selected by the Schwartz Bayesian information criterion (SIC). More importantly, we compare the log of the marginal data densities (MDD) for the models that are estimated via Bayesian methods. Based on the MDD a forecasting exercise is provided using a rolling procedure for h-steps-ahead. The GDP, CPI, FR rate and SP forecasts are estimated for the out-of-sample testing period 1997:Q1 - 2010:Q4. The forecasting investigation for the quarterly US data is performed over the one-, two-, three- and four-quarter-ahead horizon with a rolling estimation sample, based on the works of Marcellino (2004) and Brüggemann *et al.* (2008) for datasets of quarterly frequency. In particular, the models are re-estimated each quarter over the forecast horizon to update the estimate of the coefficients, before producing the quarter-ahead forecasts. Finally, in order to evaluate the models' forecast accuracy, we use the cross-model test statistic of Diebold and Mariano (1995). In particular, the modified Diebold and Mariano test (MDM) proposed by Harvey, Leybourne and Newbold (1997) is applied. The application of the MDM test is required as the DM could be seriously over-sized when the prediction horizon increases.

Firstly, we report estimation results for the log of Marginal Data Density (MDD). In particular, following Del Negro and Schorfheide (2006) we adopt the MDD as a measure of model fit, which arises naturally in the computation of posterior model odds. The prior distribution for the DSGE model parameters (θ), which are similar to the priors used by Del Negro and Schorfheide (2004), were already illustrated in Table 1. This MDD measure has two dimensions: goodness of in-sample fit on the one hand and a penalty for model complexity or degrees of freedom on the other hand. The DSGE-VAR and the DSGE-FAVAR are estimated with a different number of lags on the sample 1960:Q4 -1996:Q4. From 1997:Q1, we start our forecasting evaluation as implemented in Herbst and Schorfheide (2012). The parameter λ is chosen from a grid which is unbounded from above. In our empirical exercise, the log of the MDD is computed over a discrete interval, $\ln p(Y|\lambda, M)$. The minimum value, $\lambda_{\min} = \frac{n+k}{T}$, is model dependent and is related to the existence of a well-defined Inverse-Wishart distribution. For completeness, it is worth mentioning that $\lambda = 0$ refers to the VAR and the FAVAR model with no prior and it is not possible to compute the marginal likelihood in this particular case. Therefore, we can show the log of MDD for any value of λ larger than $\lambda_{\min}$. Importantly, $\lambda_{\min}$ depends on the degrees of freedom in the VAR or FAVAR and therefore, given estimation on the same number of available observations, $\lambda_{\min}$ for a DSGE-FAVAR will always be larger than $\lambda_{\min}$ for a DSGE-VAR[4].

Table 2 shows the main results related to the DSGE-VAR implemented using a different number of lags (from 1

[3] The Appendix contains a detailed description of all series and their corresponding transformations.

[4] For the DSGE-VAR over the sample 1960:Q4-1996:Q4, the lambda grid is given by
$$\Lambda = \left\{ \begin{array}{c} 0,\ 0.05,\ 0.08,\ 0.10,\ 0.12,\ 0.15,\ 0.20,\ 0.25, \\ 0.30,\ 0.35,\ 0.40\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9,\ 1,\ 10,\ 100 \end{array} \right\}.$$
For the DSGE-FAVAR over the sample 1960:Q4-1996:Q4, the lambda grid is given by
$$\Lambda = \left\{ \begin{array}{c} 0,\ 0.08,\ 0.1,\ 0.12,\ 0.14{,}0.15,\ 0.2,\ 0.25, \\ 0.3,\ 0{,}35,\ 0.4,\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9,\ 1,\ 10,\ 100 \end{array} \right\}.$$
In both lambda intervals, we consider the λ_{MIN} across lags from 1 to 4.

Table 2: Optimal lambda for the DSGE-VAR and DGSE-FAVAR calculated with Markov Chain Monte Carlo and Metropolis Hastings method

	λ_{MIN}	$\hat{\lambda}$	$\hat{\lambda}-\lambda_{MIN}$	$\frac{\hat{\lambda}-\lambda_{MIN}}{\lambda_{MIN}}$	$\ln p(Y\|\hat{\lambda},M)$	Bayes Factor vs M_1
DSGE-VAR(1)	0.06	0.15	0.09	1.5	-671.626	$exp[-1.468]$
DSGE-VAR(2) (M_1)	0.09	0.15	0.06	0.66	-671.493	$exp[-1.601]$
DSGE-VAR(3)	0.11	0.20	0.09	0.82	-672.247	$exp[-0.847]$
DSGE-VAR(4)	0.14	0.25	0.11	0.79	-673.094	1
	λ_{MIN}	$\hat{\lambda}$	$\hat{\lambda}-\lambda_{MIN}$	$\frac{\hat{\lambda}-\lambda_{MIN}}{\lambda_{MIN}}$	$\ln p(Y\|\hat{\lambda},M)$	Bayes Factor vs M_2
DSGE-FAVAR(1)	0.08	0.10	0.02	0.25	-649.802	$exp[15.078]$
DSGE-FAVAR(2)	0.10	0.15	0.05	0.50	-642.442	$exp[7.718]$
DSGE-FAVAR(3)	0.12	0.20	0.08	0.66	-638.746	$exp[4.022]$
DSGE-FAVAR(4) (M_2)	0.15	0.25	0.10	0.66	-634.724	1

up to 4). Each minimum λ (λ_{MIN}) is given by the features of the model (number of observations, number of endogenous variables, number of lags), and the optimal lambda ($\widehat{\lambda}$) is calculated using the Markov Chain Monte Carlo with Metropolis Hastings acceptance method (with 110,000 replications, we discard the first 10,000 ones). $\ln p(Y|M)$ is the log-MDD of the DSGE model specifications computed based on Geweke's (1999) modified harmonic mean estimator. The Bayes factor (ratio of posterior odds to prior odds), as in An and Schorfheide (2007) helps us to understand the improvement of the log-MDD of a specific model. We compare different models against the benchmark model (M) maximizing the MDD. According to Table 2, we select the DSGE-VAR with 2 lags for the full sample 1960-1996. We repeat our exercise for the DSGE-FAVAR. We select one lag for the factors and we implement - as in case of the DSGE-VAR, - the DSGE-FAVAR with a different number lags from 1 to 4. As Table 2 shows, the DSGE-FAVAR with 4 lags is chosen. In Table 3, we compare the logarithm of the MDD of the hybrid models, DSGE-VAR and DSGE-FAVAR against the DSGE, the Bayesian VAR, the VAR and the Factor Augmented VAR. The DSGE-FAVAR with 4 lags shows the maximum MDD.

Table 3: Log of the Marginal Data Density and Bayes Factor for the sample 1960:Q4-1996:Q4

	$lnp(Y\|M)$
DSGE	-693.765
DSGE-VAR(2)	-671.493
DSGE-FAVAR(4)	-634.724
BVAR(1)	-688.878
BVAR(2)	-671.757
BVAR(3)	-657.057
BVAR(4)	-645.230
VAR(1)	-673.906
VAR(2)	-672.368
VAR(3)	-672.636
VAR(4)	-672.046
FAVAR(1)	-649.580
FAVAR(2)	-646.821
FAVAR(3)	-642.736
FAVAR(4)	-639.839

Table 4 reports the RMSE for all models and variables. An exhaustive exercise was conducted on VAR and BVAR models with one to four lags based on the Schwartz Bayesian information criterion (SIC). The results provide evidence that in general four lags is the optimal number for these models. Overall, the results from RMSE are in accordance with those from the MDD estimation. Moreover, the SIC for one to four lags (MLE and QMLE estimation) was implemented in order to select the best specification for the MVSS-TVP-VAR and the heteroscedastic MVSS-TVP-MSVAR. In both cases one lag was chosen. For the GDP series the results are diverse. In particular, the MVSS-TVP-MSVAR model is the best for the one-step-ahead period, whilst the MVSS-TVP-VAR outranks the other models for two-quarters-ahead. The BVAR model provides the lowest RMSE for the three-steps ahead and only for the four-quarters-ahead the FAVAR outperforms the other models. The VAR, DSGE and DSGE-FAVAR models present similar predictive performance and on average they generate the highest forecast errors. Next, in case of the CPI variable, the DSGE model with financial frictions clearly outperforms all other models for all steps-ahead. The simple VAR and the DSGE-FAVAR outrank with a few exceptions the other model classes. The FAVAR model seems slightly better than BVAR, whilst the DSGE-VAR provides with relatively high scores for the RMSE for all quarters-ahead. The worst performers are the state-space TVP-VARs. The results for the FF rate series provide evidence of the superiority of the homoscedastic MVSS-TVP-VAR. Specifically, when comparing the RMSE scores of all model classes, the MVSS-TVP-VAR is consistently the best performer in each forecasting horizon along with the heteroscedastic version (MVSS-TVP-MSVAR). The next lowest error is produced by the DSGE model with financial frictions and the VAR model for all quarters-ahead. The DSGE-FAVAR model is better than the other for all steps-ahead, while the BVAR provides with the highest error. Finally, for the spread variable SP the simple MVSS-TVP-VAR presents the lowest RMSE for one- and two-quarters-ahead, the BVAR for the three-steps-ahead and the DSGE for the longer horizon. The next best model is the MVSS-TVP-MSVAR for the first two-quarters-ahead period as well as BVAR which outperforms all other models with the exception of the four-quarter forecast. Overall, the simple VAR shows the worst performance, whilst the DSGE-FAVAR underperforms relatively to the other models in one- and four-quarters ahead and the simple DSGE in the other two forecasting horizons[5].

Table 4: Root Mean Square Forecast Error (RMSE) for GDP, CPI, FFR and SP

	VAR	BVAR	FAVAR	DSGE	DSGE-VAR	DSGE-FAVAR	MVSS-TVP-VAR	MVSS-TVP-MSVAR
GDP								
1	0.734	0.729	0.730	0.739	0.731	0.737	0.721	0.647
2	0.743	0.734	0.735	0.755	0.738	0.747	0.687	0.690
3	0.739	0.731	0.732	0.754	0.734	0.743	0.815	0.797
4	0.737	0.737	0.735	0.754	0.741	0.745	0.860	0.837
CPI								
1	0.870	0.892	0.885	0.856	0.898	0.870	1.047	1.245
2	0.875	0.893	0.886	0.836	0.897	0.875	0.989	1.064
3	0.876	0.897	0.890	0.829	0.902	0.875	1.239	1.271
4	0.882	0.888	0.883	0.811	0.891	0.879	1.433	1.528
FFR								
1	3.753	3.992	3.954	3.694	3.977	3.821	1.427	1.521
2	3.725	4.009	3.962	3.515	3.968	3.791	1.434	1.645
3	3.845	4.127	4.085	3.594	4.105	3.909	1.788	2.001
4	4.034	4.178	4.146	3.775	4.180	4.057	2.351	2.546
SP								
1	1.261	1.216	1.221	1.243	1.221	1.250	0.790	0.946
2	1.305	1.249	1.255	1.297	1.264	1.296	1.082	1.224
3	1.308	1.253	1.259	1.304	1.261	1.298	1.491	1.660
4	1.283	1.250	1.257	1.240	1.252	1.279	2.030	2.127

[5] Interestingly, the RMSE scores are not increasing with respect to the forecast horizon (maybe with the exception of FFR). This is in accordance with the results of Del Negro and Schorfheide (2012), up to four-quarters-ahead.

Moreover, we discuss results concerning the global financial crisis period 2008 -2010 and we comparatively assess them against those of the total forecasting period. This can be considered as a robustness analysis vis-à-vis the prediction results for the total period. We consider the out-of-sample sub-period 2008:1-2010:4 and we report the RMSE for the various models over the one- to four-quarter-ahead horizon with a rolling estimation sample. For the GDP series the MVSS-TVP-MSVAR with two regimes outperforms all models for the short-term steps-ahead (one and two), possibly because it picks out the crisis period as the "high volatility" regime and the pro-crisis period as the low regime. Next, the FAVAR model is best for the three- and four-steps-ahead, whilst the BVAR provides a relatively low RMSE, very close to that of the FAVAR model. In particular, as opposed to the total sample investigation for the GDP where the results were quite diverse, for the crisis period the MVSS-TVP-VAR models present the best out-of-sample behavior in the short forecasting horizons and FAVAR in the longer predictability periods. The next best performers are the VAR, BVAR and DSGE-VAR models. In case of CPI and FFR the results are similar to the ones derived from the total out-of-sample period. The DSGE model with financial frictions - exactly as in the total period - still achieves the best scores for CPI, thus clearly outperforms all other models for all steps-ahead. The next lowest RMSE for the CPI is produced by the DSGE-FAVAR model for all quarters-ahead. For the FFR the MVSS-TVP-MSVAR achieves the best performance while in case of the total period the homoscedastic MVSS-TVP-VAR was better. This result is in accordance with the rationale of the two states/regimes for the MVSS-TVP-MSVAR, in that the heteroscedastic TVP-VAR attributes the crisis period to the "high volatility" state and hence shows a better behaviour compared to the simple TVP-VAR which proves to be better for the total period. Also, VAR shows the next best performance compared to the other models, exactly as in the total period. Finally, for the SP series the results are different from the total period. Specifically, the MVSS-TVP-MSVAR is the outranking model for the first-step-ahead and the DSGE-VAR emerges consistently (in all other quarters-ahead) as the best performer. Interestingly, the DSGE-VAR was not among the best models when the total out-of-sample period was investigated, but now when the crisis period is examined the DSGE-VAR outranks the other specifications including BVAR, yet marginally. The RMSE ratios for the global financial crisis sub-period are reported in Table 5

Table 5: Root Mean Square Forecast Error (RMSE) for GDP, CPI, FFR and SP (2008-2010 sub-period)

	VAR	BVAR	FAVAR	DSGE	DSGE-VAR	DSGE-FAVAR	MVSS-TVP-VAR	MVSS-TVP-MSVAR
GDP								
1	1.391	1.388	1.387	1.395	1.392	1.394	1.264	1.030
2	1.360	1.357	1.357	1.375	1.361	1.364	1.182	1.149
3	1.342	1.338	1.337	1.360	1.342	1.345	1.633	1.492
4	1.193	1.193	1.192	1.209	1.196	1.196	1.609	1.607
CPI								
1	1.634	1.635	1.633	1.629	1.635	1.629	2.251	2.623
2	1.701	1.702	1.699	1.660	1.702	1.696	2.106	2.283
3	1.696	1.700	1.696	1.638	1.700	1.690	2.733	2.846
4	1.694	1.691	1.688	1.606	1.692	1.688	3.246	3.444
FFR								
1	6.199	6.283	6.271	6.115	6.275	6.224	2.141	1.941
2	6.343	6.471	6.449	6.025	6.448	6.380	2.034	1.726
3	6.539	6.661	6.643	6.177	6.649	6.573	2.735	2.579
4	6.750	6.791	6.775	6.371	6.794	6.755	4.112	4.045
SP								
1	2.307	2.288	2.297	2.328	2.286	2.303	1.731	1.722
2	2.297	2.270	2.279	2.352	2.270	2.290	2.381	2.373
3	2.295	2.268	2.278	2.356	2.267	2.289	3.369	3.535
4	2.213	2.199	2.207	2.252	2.198	2.211	4.663	4.706

In the next step, the Modified Diebold-Mariano (MDM) pairwise test is employed in order to evaluate the comparative forecast accuracy. The results are reported in Tables 6, 7, 8 and 9. The Modified Diebold-Mariano test is based on the squared prediction errors. The MDM test statistics for GDP lead to a diverse and variant assessment of differential predictability, albeit many pairwise comparisons produce a statistically significant MDM score. It appears that for GDP no particular model consistently and comparatively outperforms any of the other in all steps-ahead, while the RMSE results for the MVSS-TVP-VAR and MVSS-TVP-MSVAR in one- and two-steps-ahead are not consistently verified at the 10%, 5% or 1% level in all pairwise comparisons. However, in some pairs of the BVAR model (e.g., BVAR vs. DSGE and BVAR vs. DSGE-FAVAR for three-steps-ahead etc.) differential predictability is significant at the 5% or 1% level, yet for most others this is not corroborated by the results. FAVAR which was the best performer for the four-steps-ahead produces significant results in most pairs. Considering the other pairs, the VAR model produces a significant predictability compared to DSGE and DSGE-FAVAR, the FAVAR vs. DSGE and the DSGE vs. DSGE-VAR and DSGE-FAVAR, and the two state-space time-varying VARs show significant predictability only for the first-quarter-ahead. The other examined cases in this study for GDP show weak or no differential predictability. On the contrary for the CPI series, the DSGE model with financial frictions in any pair shows a distinctively significant predictability at 1% (or 5%) in all step-ahead forecasts. This is in accordance with the RMSE results and in favor of the superior predictability of the DSGE model. In fact, most models for all forecast horizons appear to have a significant pairwise predictability at 5% or 1% level. The only exception is the pair VAR vs. DSGE-FAVAR for all steps-ahead and the pairs VAR-BVAR and VAR-FAVAR for the long forecast horizon of four quarters and FAVAR vs. DSGE-FAVAR also in four-steps-ahead. Also, MVSS-TVP-VAR and MVSS-TVP-MSVAR present statistically significant differential predictability only in case of the shorter horizon, i.e., the first-quarter-ahead. the Thus, based on the MDD, RMSE and MDM results it is evident that in case of CPI the DSGE set-up including financial frictions outperforms the other models. Similar conclusions emerge from the investigation of the pairwise forecast comparison in case of FF Rate series, in this case for the homoskedastic MVSS-TVP-VAR model. In particular, the results are indicative of a consistent outranking classification of the MVSS-TVP-VAR model vs. other models for all forecasting horizons. In fact both time-varying VAR specifications produced the lowest RMSE in a previous examination and in conjunction with the RMSE results it is shown that they consistently outperform any of the other models for all quarter-ahead forecasts, namely the differential predictability is strongly significant at 1% level. Moreover, for the FFR series it appears that almost all pairs when assessed comparatively show a strong significant differential predictability. Finally, in case of the SP variable the MDM scores depict in the majority of cases a statistically significant pairwise forecastability, albeit with many exceptions especially for longer horizons. More importantly, the MVSS-TVP-VAR model that provided with the best performance in terms of the RMSE for the first two horizons in all pairs presents a consistent assessment of differential predictability i.e., comparatively outperforms any of the other, yet only for the first-quarter-ahead and not for the second. Next, the BVAR that was the best performer for three-steps-ahead (according to RMSE) shows significant results only for some pairs and not for all of them assessed comparatively (i.e., statistical significance is indicative in case of BVAR vs VAR, FAVAR, DSGE-VAR and DSGE but not vs. MVSS-TVP-VAR or MVSS-TVP-MSVAR). Lastly, the DSGE with financial frictions that outranked any other model in the four-steps-ahead horizon produces insignificant differential predictability compared to any model. The examination of other pairs is not indicative of statistically significant predictability either. For example, VAR does not outperform the DSGE in any quarter-ahead forecast, the DSGE vs. DSGE-FAVAR pair is not producing significant scores even at the 10% level and the FAVAR is not comparatively different to the DSGE-FAVAR in almost all steps-ahead[6].

[6] The results for the differential predictability of the models based on the Modified Diebold-Mariano (MDM) are qualitatively the same for the crisis period as well.

Table 6: Pairwise forecast comparison for the GDP with the Modified Diebold-Mariano test

GDP	PERIODS			
	1	2	3	4
VAR vs BVAR	1.090	1.946	1.305	0.058
VAR vs FAVAR	0.992	1.473	1.321	2.070
VAR vs DSGE	1.414	2.025	2.528	1.743
VAR vs DSGE-VAR	0.817	1.830	1.266	1.526
VAR vs DSGE-FAVAR	4.234	3.402	2.575	2.498
VAR vs MVSS-TVP-VAR	0.154	0.523	0.473	0.833
VAR vs MVSS-TVP-MSVAR	1.031	0.436	0.417	0.627
BVAR vs FAVAR	1.174	0.948	0.209	1.714
BVAR vs DSGE	2.298	3.212	2.265	1.778
BVAR vs DSGE-VAR	1.307	1.621	1.166	4.569
BVAR vs DSGE-FAVAR	1.663	2.373	1.577	2.835
BVAR vs MVSS-TVP-VAR	0.099	0.447	0.521	0.832
BVAR vs MVSS-TVP-MSVAR	0.979	0.366	1.361	0.626
FAVAR vs DSGE	2.273	3.222	2.254	2.119
FAVAR vs DSGE-VAR	0.974	1.425	1.098	2.844
FAVAR vs DSGE-FAVAR	1.599	2.389	1.591	2.573
FAVAR vs MVSS-TVP-VAR	0.106	0.452	0.521	0.846
FAVAR vs MVSS-TVP-MSVAR	0.988	0.370	1.361	0.639
DSGE vs DSGE-VAR	2.080	2.757	2.541	1.493
DSGE vs DSGE-FAVAR	0.625	1.279	1.964	0.862
DSGE vs MVSS-TVP-VAR	0.219	0.628	0.379	0.712
DSGE vs MVSS-TVP-MSVAR	1.093	0.531	1.325	0.518
DSGE-VAR vs DSGE-FAVAR	1.696	2.603	1.682	1.626
DSGE-VAR vs MVSS-TVP-VAR	0.126	0.479	0.505	0.807
DSGE-VAR vs MVSS-TVP-MSVAR	1.003	0.396	1.359	0.604
DSGE-FAVAR vs MVSS-TVP-VAR	0.190	0.562	0.447	0.779
DSGE-FAVAR vs MVSS-TVP-MSVAR	1.064	0.472	0.387	0.579
MVSS-TVP-VAR vs MVSS-TVP-MSVAR	1.723	0.129	0.386	0.568

Notes: The Modified Diebold-Mariano (1995) test is based on squared prediction errors and is distributed as Student with (T-1) degrees of freedom, where T the out-of-sample sample period. The reported numbers are the t-scores.

6 Conclusions and policy implications

The financial crisis revealed that the workhorse of contemporary DSGE modelling is not fully ready to deal with the issue of credit frictions and financial stability, which play an important role in modern business cycle theory. Recently there have been some efforts to model financial factors mainly by Goodfriend and McCallum (2007), Christiano *et al.* (2010), Curdía and Woodford (2010), Gerali *et al.* (2010), Gertler and Kiyotaki (2010) and Gertler and Karadi (2011). These constitute some promising attempts to study the effects of financial intermediation on business cycle fluctuations and policy design. In particular, DSGE models with financial frictions could deal with many issues, such as the effect of financial shocks on real variables, how optimal monetary policy is affected by the presence of financial frictions and how macroprudential policies impact capital requirements. Moreover, new macroeconomic research is drawn to the application of Bayesian statistics because DSGE models are often seen as abstractions of actual economies. In this study we developed a DSGE model by augmenting the small-scale model with financial frictions, following Christiano et al. (2009) and Del Negro and Schorfheide (2012). In this way we investigated the response of the monetary policy to increases in interest rate spreads. We included a comparative evaluation of the out-of-sample predictive performance of many different specifications of estimated DSGE models and various classes of VAR models, using datasets from the US economy. Simple and hybrid DSGE models were implemented, such as DSGE-VARs and Factor Augmented DGSEs (DSGE-FAVAR), and tested against standard VARs, Bayesian VARs and Factor Augmented VARs (FAVAR). Moreover, we introduced two novel specifications for time-varying parameter autoregressive models (TVP-VARs) in an attempt to overcome restrictive stationarity assumptions, to incorporate uncertainty due to future asymmetric random shocks and more importantly to account for inherent nonlinearities and adaptive relationships of the economy that cannot be captured by the atheoretic VAR family models as well as

Table 7: Pairwise forecast comparison for the CPI with the Modified Diebold-Mariano test

CPI	PERIODS			
	1	2	3	4
VAR vs BVAR	4.659	2.910	2.702	1.673
VAR vs FAVAR	4.090	2.245	2.282	0.488
VAR vs DSGE	1.943	3.309	2.866	2.856
VAR vs DSGE-VAR	4.906	3.034	2.803	1.943
VAR vs DSGE-FAVAR	0.031	0.364	0.461	1.584
VAR vs MVSS-TVP-VAR	1.390	0.597	0.873	0.908
VAR vs MVSS-TVP-MSVAR	1.237	0.909	0.905	0.984
BVAR vs FAVAR	5.375	4.662	4.182	4.056
BVAR vs DSGE	3.596	3.799	3.239	3.011
BVAR vs DSGE-VAR	5.554	2.694	3.136	2.173
BVAR vs DSGE-FAVAR	5.118	3.145	3.024	2.319
BVAR vs MVSS-TVP-VAR	1.332	0.503	0.826	0.899
BVAR vs MVSS-TVP-MSVAR	1.176	0.825	0.861	0.977
FAVAR vs DSGE	3.294	3.666	3.130	2.949
FAVAR vs DSGE-VAR	5.615	4.052	3.703	3.095
FAVAR vs DSGE-FAVAR	4.793	2.558	2.658	1.438
FAVAR vs MVSS-TVP-VAR	1.350	0.539	0.840	0.904
FAVAR vs MVSS-TVP-MSVAR	1.195	0.857	0.874	0.982
DSGE vs DSGE-VAR	3.847	3.812	3.298	3.045
DSGE vs DSGE-FAVAR	1.954	3.456	2.882	2.886
DSGE vs MVSS-TVP-VAR	1.427	0.796	0.960	0.990
DSGE vs MVSS-TVP-MSVAR	1.276	1.073	0.988	1.059
DSGE-VAR vs DSGE-FAVAR	5.313	3.271	3.090	2.393
DSGE-VAR vs MVSS-TVP-VAR	1.315	0.484	0.814	0.893
DSGE-VAR vs MVSS-TVP-MSVAR	1.158	0.808	0.850	0.973
DSGE-FAVAR vs MVSS-TVP-VAR	1.389	0.599	0.872	0.910
DSGE-FAVAR vs MVSS-TVP-MSVAR	1.237	0.910	0.905	0.987
MVSS-TVP-VAR vs MVSS-TVP-MSVAR	0.800	1.310	0.463	0.286

Notes: As in Table 6

Table 8: Pairwise forecast comparison for the FF rate with the Modified Diebold-Mariano test

FFR	PERIODS			
	1	2	3	4
VAR vs BVAR	6.541	3.564	2.916	2.176
VAR vs FAVAR	6.485	3.464	2.863	2.022
VAR vs DSGE	4.972	4.216	3.558	2.663
VAR vs DSGE-VAR	6.407	3.641	2.924	2.324
VAR vs DSGE-FAVAR	6.427	4.059	3.147	2.029
VAR vs MVSS-TVP-VAR	5.834	3.368	2.662	2.566
VAR vs MVSS-TVP-MSVAR	5.405	2.943	2.302	2.141
BVAR vs FAVAR	6.802	4.116	3.229	2.842
BVAR vs DSGE	7.658	4.847	3.952	3.098
BVAR vs DSGE-VAR	4.450	2.651	2.045	0.387
BVAR vs DSGE-FAVAR	6.527	3.389	2.815	2.103
BVAR vs MVSS-TVP-VAR	6.324	3.681	2.932	2.658
BVAR vs MVSS-TVP-MSVAR	5.860	3.244	2.562	2.240
FAVAR vs DSGE	7.752	4.871	3.971	3.066
FAVAR vs DSGE-VAR	4.489	0.758	2.063	3.508
FAVAR vs DSGE-FAVAR	6.424	3.215	2.721	1.886
FAVAR vs MVSS-TVP-VAR	6.250	3.630	2.894	2.639
FAVAR vs MVSS-TVP-MSVAR	5.790	3.194	2.525	2.217
DSGE vs DSGE-VAR	7.419	4.910	3.948	3.136
DSGE vs DSGE-FAVAR	7.305	4.754	3.866	2.730
DSGE vs MVSS-TVP-VAR	5.753	3.205	2.496	2.508
DSGE vs MVSS-TVP-MSVAR	5.312	2.761	2.117	2.006
DSGE-VAR vs DSGE-FAVAR	6.374	3.475	2.832	2.280
DSGE-VAR vs MVSS-TVP-VAR	6.278	3.642	2.907	2.665
DSGE-VAR vs MVSS-TVP-MSVAR	5.820	3.209	2.541	2.248
DSGE-FAVAR vs MVSS-TVP-VAR	5.972	3.445	2.723	2.575
DSGE-FAVAR vs MVSS-TVP-MSVAR	5.534	3.018	2.363	2.155
MVSS-TVP-VAR vs MVSS-TVP-MSVAR	0.737	1.015	0.882	2.393

Notes: As in Table 6

Table 9: Pairwise forecast comparison for the SP with the Modified Diebold-Mariano test

SP	PERIODS			
	1	2	3	4
VAR vs BVAR	6.645	4.102	3.177	2.875
VAR vs FAVAR	6.289	3.884	2.991	2.596
VAR vs DSGE	3.029	0.474	0.143	1.125
VAR vs DSGE-VAR	6.990	4.204	3.310	2.872
VAR vs DSGE-FAVAR	7.070	4.571	3.722	1.845
VAR vs MVSS-TVP-VAR	4.127	0.712	0.313	0.791
VAR vs MVSS-TVP-MSVAR	2.787	0.429	0.687	0.974
BVAR vs FAVAR	4.546	3.107	2.580	2.443
BVAR vs DSGE	4.131	2.683	2.053	0.323
BVAR vs DSGE-VAR	3.522	3.148	2.004	1.187
BVAR vs DSGE-FAVAR	6.496	3.963	3.052	2.731
BVAR vs MVSS-TVP-VAR	3.810	0.532	0.401	0.821
BVAR vs MVSS-TVP-MSVAR	2.405	0.133	0.783	1.006
FAVAR vs DSGE	3.981	2.624	1.991	0.562
FAVAR vs DSGE-VAR	0.357	1.816	0.398	1.508
FAVAR vs DSGE-FAVAR	6.007	3.658	2.801	2.333
FAVAR vs MVSS-TVP-VAR	3.828	0.551	0.392	0.816
FAVAR vs MVSS-TVP-MSVAR	2.433	0.165	0.774	1.000
DSGE vs DSGE-VAR	3.450	1.761	1.705	0.372
DSGE vs DSGE-FAVAR	1.351	0.065	0.293	1.034
DSGE vs MVSS-TVP-VAR	3.900	0.684	0.320	0.837
DSGE vs MVSS-TVP-MSVAR	2.557	0.384	0.701	1.026
DSGE-VAR vs DSGE-FAVAR	6.887	4.080	3.196	2.684
DSGE-VAR vs MVSS-TVP-VAR	3.852	0.583	0.389	0.820
DSGE-VAR vs MVSS-TVP-MSVAR	2.453	0.216	0.769	1.004
DSGE-FAVAR vs MVSS-TVP-VAR	4.046	0.684	0.329	0.795
DSGE-FAVAR vs MVSS-TVP-MSVAR	2.692	0.383	0.704	0.978
MVSS-TVP-VAR vs MVSS-TVP-MSVAR	1.969	1.268	1.344	0.246

Notes: As in Table 6

by the micro-founded DSGE models according to the Lucas critique. In this paper, a novel time-varying multivariate state-space estimation method for TVP-VAR processes is proposed both for homoskedastic and heteroskedastic error structures. For the TVP-VAR models, the parameters are estimated using a multivariate specification of the standard Kalman set-up (Harvey, 1990; Bekiros and Paccagnini, 2013) with extended quasi-optimal filtering in particular for the TVP-VAR with Markov-switching heteroskedasticity. The likelihood estimation of the TVP-VAR is performed with a suitable multivariate extension of the Kim and Nelson (1999a, 1999b) method. The results were assessed with the use of Bayesian marginal data density and the Quasi-MLE method and evaluated with the root mean squared forecast error. The modified Diebold and Mariano test (MDM) proposed by Harvey, Leybourne and Newbold (1997) was also employed to measure comparatively the differential forecastability taking into account the serious over-sizedness of the simple DM test when the prediction horizon increases.

The best forecasting performance for the CPI series was consistently produced by the DSGE model with financial frictions for all forecast horizons. In case of the FFR variable the homoscedastic TVP-VAR model outperforms all other models for all steps-ahead, being equal in forecasting performance with the TVP-VAR with Markov-Switching heteroskedasticity. For the GDP series the results were diverse. In particular, the heteroskedastic TVP-VAR model is the best for the one-step-ahead period, whilst the standard TVP-VAR outranks the other models for two-quarters-ahead. Next, the BVAR model provides the lowest RMSE for the three-steps ahead and only for the four-quarter-ahead the FAVAR outperforms the other models. Similarly, mixed results are produced for the SP series: the homoskedastic TVP-VAR presents the lowest RMSE for one- and two-quarters-ahead, the BVAR for the three-steps-ahead and the DSGE with financial frictions for the longer horizon.

This work is open to several extensions. From the modelling point of view it should be relevant to focus on the interaction of frictions both at firms level and in the banking sector, in order to examine the transmission mechanism of the shocks and the accelerator/attenuator effects in line with the recent contribution by Rannenberg (2012). Moreover, the recent episodes of financial turmoil have reinforced the idea that the business cycle might be, to some extent, the result of changes in agents' expectations. Empirical evidence reported in Mankiw and Reis (2002), Orphanides and Williams (2008) and Lansing (2009) is in favour of the fact that expectations concerning inflation and other economic variables display systematic mistakes that increase the stochastic volatility in the economy. Further research is needed to find macro-financial micro-founded DSGE models as well as adaptive TVP-VARs, which are able to reflect reality better. For instance, featuring a more explicit modeling of financial intermediation or introducing occasionally binding constraints seem interesting avenues.

References

Adolfson M, Stefan L, Jesper L, Mattias V (2008) Evaluating an Estimated New Keynesian Small Open Economy Model, Journal of Economic Dynamics and Control, 32(8), 2690-2721

Altug S (1989) Time-to-Build and Aggregate Fluctuations: Some New Evidence, International Economic Review, 30(4), 889-920

Bai J., Ng S. (2000) Determining the Number of Factors in Approximate Factor Models, Econometrica, 70.

Bekiros S., Paccagnini A., (2013) On the predictability of time-varying VAR and DSGE models, Empirical Economics, 45 (1), 635-664

Benati L, Surico P (2009) VAR Analysis and the Great Moderation, American Economic Review, 99(4), 1636-52

Bernanke B S., Boivin J. (2003) Monetary Policy in a Data-Rich Environment, Journal of Monetary Economics 50 (3), 525-546.

Bernanke B S., Boivin J. and Eliasz P. (2005) Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach, The Quarterly Journal of Economics, MIT Press, vol. 120(1), pages 387-422, January.

Bernanke B, Gertler M, Girlstichr S, 1999. The financial accelerator in a quantitative business cycle framework in: J. B. Taylor & M. Woodford (ed.), Handbook of Macroeconomics, edition 1, volume 1, chapter 21, pp 1341-1393

Boivin J, Giannoni MP (2006) DSGE Models in a Data-Rich Environment, NBER Working Papers 12772

Boivin J, Giannoni MP and Mihov I. (2009) Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data, American Economic Review, American Economic Association, vol. 99(1), 350-84

Branch W., Evans, G., 2007. Model Uncertainty and Endogenous Volatility, Review of Economic Dynamics, Elsevier for the Society for Economic Dynamics, vol. 10(2), pages 207-237

Brzoza-Brzezina M, M Kolasa, 2012. Bayesian evaluation of DSGE models with financial frictions, National Bank of Poland Working Paper No. 109.

Brüggemann R, H. Lütkepohl and M. Marcellino (2008) Forecasting euro area variables with German pre-EMU data, Journal of Forecasting, 27(6), 465–481

Bullard J, Suda J, 2011. The stability of macroeconomic systems with Bayesian learners, Banque de France Working Paper No. 332.

Canova F (1994) Statistical Inference in Calibrated Models, Journal of Applied Econometrics, 9, 123-144

Canova F, Gambetti L (2004) On the Time Variations of US Monetary Policy: Who is Right?, Money Macro and Finance (MMF) Research Group Conference.

Christensen I, Dib A, 2008. The financial accelerator in an estimated New Keynesian model, Review of Economic Dynamics, 11,1:155-178.

Christiano L, Motto R, Rostagno M, 2010. Financial Factors in Economics Fluctuations, Working Paper Series 1192, European Central Bank.

Christiano, L. J., M. Eichenbaum, C. Evans (2005) Nominal Rigidities and the Dynamic effects of a Shock to Monetary Policy, Journal of Political Economy 113, 1-45.

Christoffel K, Günter C, Anders W (2008) The New Area-Wide Model of the Euro Area - A Micro-Founded Open-Economy Model for Forecasting and Policy Analysis, European Central Bank Working Paper Series n. 944

Chudik, A. and M. H. Pesaran (2009) Infinite Dimensional VARs and Factor Models, Working Paper Series 998, European Central Bank.

Clarida R, Galí J, Gertler M (2000) Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory, The Quarterly Journal of Economics, 115(1), 147-180

Cogley TF, Sargent TJ (2005) Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S, Review of Economic Dynamics, 8(2), 262-302

Cochrane JH, 2007. Identification with Taylor Rules: A Critical Review, NBER Working Papers 13410, National Bureau of Economic Research, Inc.

Consolo A, Favero CA and Paccagnini A (2009) On the Statistical Identification of DSGE Models, Journal of Econometrics, 150, 99-115.

Curdia V, Woodford M. 2010. Credit Spreads and Monetary Policy , Journal of Money, Credit and Banking, Blackwell Publishing, vol. 42(s1), pages 3-35, 09.

Dahlhaus R (1997) Fitting time series models to nonstationary processes, The Annals of Statistics, 25, 1-37

Dahlhaus R (2000) A likelihood approximation for locally stationary processes, The Annals of Statistics, 28, 1762-1794

DeJong D, Ingram B, Whiteman C (1996) A Bayesian Approach to Calibration, Journal of Business Economics and Statistics, 14, 1-9

Del Negro M, Schorfheide F (2004) Priors from General equilibrium Models for VARs, International Economic Review, 45, 643-673

Del Negro M, Schorfheide F (2006) How Good Is What You've Got? DGSE-VAR as a Toolkit for Evaluating DSGE Models, Federal Reserve Bank of Atlanta Economic Review, 2nd Quarter, 21-37.

Del Negro M, Schorfheide F, Smets F, Wouters R (2007) On the Fit of New-Keynesisan Models, Journal of Business, Economics and Statistics, 25(2), 124-162

Del Negro M, Schorfheide F, (2012) DSGE Model-Based Forecasting, prepared for Handbook of Economic Forecasting, Volume 2.

Diebold FX, Mariano RS (1995) Comparing predictive accuracy, Journal of Business and Economic Statistics, 13(3), 253-263

Doan T, Litterman R, Sims C (1984) Forecasting and Conditional Projections Using Realistic Prior Distributions, Econometric Reviews, 3, 1-100

Dua P, Ray SC (1995) A BVAR model for the Connecticut Economy, Journal of Forecasting, 14(3), 167-180

Durbin J, Koopman SJ (2002) Time Series Analysis by State Space Methods, Oxford University Press

Fagan G., Henry J. and Mestre R. (2001) An area-wide model (AWM) for the euro area, Working Paper Series 42, European Central Bank.

Fernández-Villaverde J, Rubio-Ramírez, J, 2007. How Structural Are Structural Parameters?, NBER Working Papers 13166, National Bureau of Economic Research, Inc.

Fernández-Villaverde J, Guerron-Quintana PA, Rubio-Ramírez JF, 2010. Fortune or Virtue: Time-Variant Volatilities Versus Parameter Drifting in U.S. Data, NBER Working Papers 15928, National Bureau of Economic Research, Inc.

Fernández-Villaverde J, Guerron-Quintana PA, Rubio-Ramírez JF, 2010. The New Macroeconometrics: A Bayesian Approach, in A. O'Hagan and M. West, eds., Handbook of Applied Bayesian Analysis, Oxford: Oxford University Press.

Fernandez-Villaverde J (2009) The Econometrics of DSGE Models, NBER Working Paper 14677

Fernández-de-Córdoba G, Torres JL (2010) Forecasting the Spanish Economy with a DSGE Model: An Augmented VAR Approach, Journal of the Spanish Economic Association. Forthcoming

Forni M, Reichlin L. (1996): Dynamic Common Factors in Large Cross-Sections, Empirical Economics, 21, 27-42.

Forni M, Reichlin L. (1998): Let's Get Real: A Dynamic Factor Analytical Approach to Disaggregated Business Cycle, Review of Economic Studies, 65, 453-474.

Forni M, Hallin M, Lippi M and Reichlin L. (1999): Reference Cycles: The NBER Methodology Revisited, mimeo.

Forni M, Hallin M, Lippi M and Reichlin L. (2000): The Generalized Dynamic-Factor Model: Identification And Estimation, The Review of Economic and Statistics, MIT Press, vol. 82(4), pages 540-554, November.

Gelain P, Lansing K, Mendocino C, 2012. House Prices, Credit Growth, and Excess Volatility: Implications for Monetary and Macroprudential Policy, Federal Reserve Bank of San Francisco Working Paper No. 11.

Gerali A, Neri S, Sessa L, Signoretti F, 2010. Credit and Banking in a DSGE Model of the Euro Area. Journal of Money, Credit and Banking, 42:107-141.

Geweke J (1999) Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication, Econometric Reviews 18(1), 1-126

Ghent A (2009) Comparing DSGE-VAR Forecasting Models: How Big are the Differences?, Journal of Economic Dynamics and Control, 2009, 33(4), 864-882

Gertler M, Karadi, P, 2011. A Model of Unconventional Monetary Policy, Journal of Monetary Economics, Elsevier, vol. 58(1), pages 17-34.

Gertler M, Kiyotaki N, 2010. Financial Intermediation and Credit Policy in Business Cycle Analysis, Discussion paper.

Gilchrist S, Zakrajšek E, 2011. Monetary policy and credit supply shocks. IMF Economic Review, 59(2):195_232.

Goodfriend M, McCallum BT, 2007. Banking and interest rates in monetary policy analysis: A quantitative exploration, Journal of Monetary Economics, Elsevier, vol. 54(5), pages 1480-1507

Hamilton, J . (1988), Rational-Expectation Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates, Journal of Economic Dynamics and Control, 12, 385-423.

Harvey AC (1990) Forecasting, structural time series and the Kalman filter, Cambridge University Press

Harvey AC (1986) Analysis and generalization of a multivariate exponential smoothing model, Management Science, 32, 374-380

Harvey, D., S. Leybourne and P. Newbold (1997) Testing the equality of prediction mean squared errors. International Journal of Forecasting 13, 281–291.

Ingram B, Whiteman C (1994) Supplanting the Minnesota Prior - Forecasting Macroeconomics Time Series using Real Business Cycle Model Priors, Journal of Monetary Economics, 34, 497-510

Iacoviello M, 2005. House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle, American Economic Review, American Economic Association, vol. 95(3), pages 739-764, June.

Ireland P (2004) A Method for Taking Models to the Data, Journal of Economic Dynamics and Control, 28, 1205-1226

Justiniano A, Primiceri GE (2008) The Time-Varying Volatility of Macro-economic Fluctuations, American Economic Review, 98(3), 604-41

Kadiyala KR, Karlsson S (1997) Numerical Methods for Estimation and Inference in Bayesian VAR-Models, Journal of Applied Econometrics, 12(2), 99-132

Kalman RE (1960) A New Approach to Linear Filtering and Prediction. Journal of Basic Engineering, Transactions of the ASME Series D 82, 35-45

Kalman RE (1963) New Methods in Wiener Filtering Theory. In: Bogdanoff J.L., Kozin F. (eds.) Proceedings of the First Symposium of Engineering Applications of Random Function Theory and Probability, Wiley, New York, 270-388

Kim C-J (1993) Unobserved-Component Time Series Models with Markov-Switching Heteroscedasticity: Changes in Regime and the Link between Inflation Rates and Inflation Uncertainty, Journal of Business & Economic Statistics, 11 (3), 341-349

Kim C-J, Nelson CR (1999) Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle, The Review of Economics and Statistics, MIT Press, 81(4), 608-616

Kim C-J, Nelson CR (1999) State space models with Regime switching, MIT Press

King RG (2000) The New IS-LM Model: Language, Logic, and Limits, Federal Reserve Bank of Richmond Economic Quarterly, 86, 45-103

Kiyotaki N, Moore J, 1997. Credit Cycles, Journal of Political Economy, 105, pp. 211-248.

Kolasa M, Rubaszek M, Skrzypczynski P (2009) Putting the New Keynesian DSGE Model to the Real-Time Forecasting Test, European Central Bank Working Paper Series n. 1110

Lamoureux,C . G., and Lastrapes,W. (1990), Persistence in Variance, Structural Change, and the GARCH Model, Journal of Business & Economic Statistics, 8, 225- 234.

Lansing K, 2009. Time Varying U.S. Inflation Dynamics and the New Keynesian Phillips Curve, Review of Economic Dynamics, Elsevier for the Society for Economic Dynamics, vol. 12(2), pages 304-326.

Lastrapes, W. D. (1989), Exchange Rate Volatility and U.S. Monetary Policy: An ARCH Application, Journal of Money, Credit, and Banking, 21, 66-77.

Litterman RB (1981) A Bayesian Procedure for Forecasting with Vector Autoregressions, Working Paper, Federal Reserve Bank of Minneapolis

Litterman RB (1986) Forecasting with Bayesian Vector Autoregressions: Five Years of Experience, Journal of Business and Statistics 4(1), 25–38

Lubik TA, Schorfheide F (2004) Testing for Indeterminacy: An Application to U.S. Monetary Policy, American Economic Review, 94(1), 190-217

Lucas RE (1976) Econometric policy evaluation: a critique, Carnegie Rochester Conference Series on Public Policy, 1, 19-46

Marcellino, M. (2004) Forecasting EMU macroeconomic variables, International Journal of Forecasting, 20, 359–372

Marcellino, M., J. H. Stock and M.W. Watson (2003) Macroeconomic forecasting in the Euro area: Country specific versus area-wide information, European Economic Review, 47(1), 1–18

Mankiw G, Reis R, 2002. Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve, Quarterly Journal of Economics (2002) 1295–1328.

Mavroeidis S, 2005. Identification issues in forward-looking models estimated by GMM, with an application to the Phillips curve, Journal of Money Credit and Banking, 37, pp. 421-448.

McConnell MM, Perez-Quiros G (2000) Output Fluctuations in the United States: What Has Changed since the Early 1980's? American Economic Review, 90(5), 1464-1476

Orphanides A, Williams JC, 2008. Learning, expectations formation, and the pitfalls of optimal control monetary policy, Journal of Monetary Economics, Vol. 55.

Prado G.H, West M (2001) Bayesian time-varying autoregressions: Theory, methods and applications. Resenhas Journal, 4, 405-422

Primiceri GE (2005) Time Varying Structural Vector Autoregressions and Monetary Policy, The Review of Economic Studies, 72, 821-852

Queijo von Heideken V, 2009. How Important are Financial Frictions in the United States and the Euro Area? Scandinavian Journal of Economics, 111(3):567-596.

Rannenberg A, 2012. Asymmetric information in credit markets, bank leverage cycles and macroeconomic dynamics. National Bank of Belgium Working Paper Research, 224.

Rubaszek M, Skrzypczynski P (2008) On the Forecasting Perfomance of a Small-Scale DSGE Model, International Journal of Forecasting, 24, 498-512

Sargent T (1989) Two Models of Measurements and the Investment Accelerator, Journal of Political Economy, 97(2), 251-287

Schorfheide F (2010) Estimation and Evaluation of DSGE Models: Progress and Challenges, University of Pennsylvania

Schorfheide F, Sill K, Kryshko M (2010) DSGE Model-Based Forecasting of Non-Modelled Variables, International Journal of Forecasting, 26(2), 348-373

Sims CA (1980) Macroeconomics and reality, Econometrica, 48(1), 1-48

Sims, CA (2002) Solving Linear Rational Expectations Models, Computational Economics, 20 (1-2), 1-20

Sims CA, Zha T (1998) Bayesian Methods for Dynamic Multivariate Models, International Economic Review, 39, 949-968

Sims CA, Zha T (2006) Were There Regime Switches in U.S. Monetary Policy?, American Economic Review, 96(1), 54-81

Smets F, Wouters R (2003) An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area, Journal of the European Economic Association, 1, 1123-1175

Smets F, Wouters R (2004) Forecasting with a Bayesian DSGE model: an application to the Euro area, Working Paper No. 389, European Central Bank, Frankfurt

Spencer DE (1993) Developing a Bayesian Vector Autoregression Forecasting Model, International Journal of Forecasting 9(3), 407–421

Stock JH, Watson WM (2001) Vector autoregressions, Journal of Economic Perspectives, 15, 101-115

Stock JH, Watson WM (2002) Macroeconomic Forecasting Using Diffusion Indexe , Journal of Business Economics and Statistics, XX:II, 147-162.

Theil H, Goldberg AS (1961) On Pure and Mixed Estimation in Economics, International Economic Review, 2, 65-78

Todd RM (1984) Improving Economic Forecasting with Bayesian Vector Autoregression, Quarterly Review, Federal Reserve Bank of Minneapolis

Townsend R, 1979. Optimal Contracts and Competitive Markets with Costly State Verification, Journal of Economic Theory vol. 21, pp-265-293.

Woodford M (2003) Interest and Prices, Princeton University Press

Warne A, Coenen G, Christoffel K, 2012. Forecasting with DSGE-VAR Models, Manuscript, European Central Bank.

7 Appendix

The source of the data is the Federal Reserve Economic Data - Federal Reserve Bank of Saint Louis (http://research.stlouisfed.org/fred2/). In order to construct the FAVAR we extract factors from a balanced panel of 109 monthly and quarterly macroeconomic and financial time series, following the dataset built by Stock and Watson (2002). The dataset involves several measures of industrial production, interest rates, various price indices, employment and other important macroeconomic and also financial variables. In the following Table, the first column has the series number, the second the series acronym, the third the series description, the fourth the transformation codes and the fifth column denotes a slow-moving variable with 1 and a fast-moving one with 0. The transformed series are tested using the Box-Jenkins procedure and the Dickey-Fuller test. Following Bernanke *et al.* (2005), the transformation codes are as follows: 1 - no transformation; 2 - first difference; 4 - logarithm; 5 - first difference of logarithm; 6 - second difference; 7 - second difference of logarithm.

Date	Long Description	Tcode	SlowCode
PAYEMS	Total Nonfarm Payrolls: All Employees	5	1
DSPIC96	Real Disposable Personal Income	5	1
NAPM	ISM Manufacturing: PMI Composite Index	1	1
UNRATE	Civilian Unemployment Rate	1	1
INDPRO	Industrial Production Index (Index 2007=100)	5	1
PCEPI	Personal Consumption Expenditures: Chain-type Price Index (Index 2005=100)	5	1
PPIACO	Producer Price Index: All Commodities (Index 1982=100)	5	1
FEDFUNDS	Effective Federal Funds Rate	1	0
IPDCONGD	Industrial Production: Durable Consumer Goods (Index 2007=100)	5	1
IPBUSEQ	Industrial Production: Business Equipment (Index 2007=100)	5	1
IPMAT	Industrial Production: Materials (Index 2007=100)	5	1
IPCONGD	Industrial Production: Consumer Goods (Index 2007=100)	5	1
IPNCONGD	Industrial Production: Nondurable Consumer Goods (Index 2007=100)	5	1
IPFINAL	Industrial Production: Final Products (Market Group) (Index 2007=100)	5	1
UNEMPLOY	Unemployed	5	1
EMRATIO	Civilian Employment-Population Ratio (%)	1	1
CE16OV	Civilian Employment	5	1
CLF16OV	Civilian Labor Force	5	1
CIVPART	Civilian Participation Rate (%)	1	1
UEMP27OV	Civilians Unemployed for 27 Weeks and Over	5	1
UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	5	1
UEMP15OV	Civilians Unemployed - 15 Weeks & Over	5	1
UEMP15T26	Civilians Unemployed for 15-26 Weeks	5	1
UEMP5TO14	Civilians Unemployed for 5-14 Weeks	5	1
MANEMP	Employees on Nonfarm Payrolls: Manufacturing	5	1
USPRIV	All Employees: Total Private Industries	5	1
USCONS	All Employees: Construction	5	1
USFIRE	All Employees: Financial Activities	5	1
USTRADE	All Employees: Retail Trade	5	1
DMANEMP	All Employees: Durable Goods Manufacturing	5	1
USGOOD	All Employees: Goods-Producing Industries	5	1
USEHS	All Employees: Education & Health Services	5	1
USLAH	All Employees: Leisure & Hospitality	5	1
SRVPRD	All Employees: Service-Providing Industries	5	1
USINFO	All Employees: Information Services	5	1
USPBS	All Employees: Professional & Business Services	5	1
USTPU	All Employees: Trade, Transportation & Utilities	5	1
NDMANEMP	All Employees: Nondurable Goods Manufacturing	5	1
USMINE	All Employees: Natural Resources & Mining	5	1
USWTRADE	All Employees: Wholesale Trade	5	1
USSERV	All Employees: Other Services	5	1
AHEMAN	Average Hourly Earnings: Manufacturing	5	1
AHECONS	Average Hourly Earnings: Construction (NSA)	5	1
PPIIDC	Producer Price Index: Industrial Commodities (NSA)	5	1

PPIFGS	Producer Price Index: Finished Goods (Index 1982=100)	5	1
PPICPE	Producer Price Index: Finished Goods: Capital Equipment (Index 1982=100)	5	1
PPICRM	Producer Price Index: Crude Materials for Further Processing (Index 1982=100)	5	1
PPIITM	Producer Price Index: Intermediate Materials: Supplies & Components (Index 1982=100)	5	1
PPIENG	Producer Price Index: Fuels & Related Products & Power (Index 1982=100)	5	1
PPIFCG	Producer Price Index: Finished Consumer Goods (Index 1982=100)	5	1
PFCGEF	Producer Price Index: Finished Consumer Goods Excluding Foods (Index 1982=100)	5	1
CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items (Index 1982=100)	5	1
CPIAUCNS	Consumer Price Index for All Urban Consumers: All Items (Index 1982-84=100)	5	1
CPILFESL	Consumer Price Index for All Urban Consumers: All Items Less Food & Energy (Index 1982-84=100)	5	1
CPILFENS	Consumer Price Index for All Urban Consumers: All Items Less Food & Energy (NSA Index 1982=100)	5	1
CPIUFDNS	Consumer Price Index for All Urban Consumers: Food (NSA Index 1982=100)	5	1
CPIENGNS	Consumer Price Index for All Urban Consumers: Energy (NSA Index 1982=100)	5	1
CPIENGSL	Consumer Price Index for All Urban Consumers: Energy (Index 1982-1984=100)	5	1
CPILEGSL	Consumer Price Index for All Urban Consumers: All Items Less Energy (Index 1982-1984=100)	5	1
CPIMEDSL	Consumer Price Index for All Urban Consumers: Medical Care (Index 1982-1984=100)	5	1
PPIFCF	Producer Price Index: Finished Consumer Foods (Index 1982=100)	5	1
AAA	Moody's Seasoned Aaa Corporate Bond Yield	1	0
BAA	Moody's Seasoned Baa Corporate Bond Yield	1	0
M2SL	M2 Money Stock	6	0
M2NS	M2 Money Stock (NSA)	6	0
M1NS	M1 Money Stock (NSA)	6	0
M3SL	M3 Money Stock (DISCONTINUED SERIES)	6	0
GS5	5-Year Treasury Constant Maturity Rate	1	0
GS10	10-Year Treasury Constant Maturity Rate	1	0
GS1	1-Year Treasury Constant Maturity Rate	1	0
GS3	3-Year Treasury Constant Maturity Rate	1	0
TB3MS	3-Month Treasury Bill: Secondary Market Rate	1	0
TB6MS	6-Month Treasury Bill: Secondary Market Rate	1	0
HOUST	Housing Starts: Total: New Privately Owned Housing Units Started	5	0
PERMIT	New Private Housing Units Authorized by Building Permits	5	0
HOUSTMW	Housing Starts in Midwest Census Region	5	0
HOUSTW	Housing Starts in West Census Region	5	0
HOUSTNE	Housing Starts in Northeast Census Region	5	0
HOUSTS	Housing Starts in South Census Region	5	0
PERMITS	New Private Housing Units Authorized by Building Permits - South	5	0
PERMITMW	New Private Housing Units Authorized by Building Permits - Midwest	5	0
PERMITW	New Private Housing Units Authorized by Building Permits - West	5	0
PERMITNE	New Private Housing Units Authorized by Building Permits - Northeast	5	0
PDI	Personal Dividend Income	5	0
SPREAD1	3mo-FYFF	1	0
SPREAD2	6mo-FYFF	1	0
SPREAD3	1yr-FYFF	1	0
SPREAD4	2yr-FYFF	1	0
SPREAD5	3yr-FYFF	1	0
SPREAD6	5yr-FYFF	1	0
SPREAD7	7yr-FYFF	1	0
SPREAD8	10yr-FYFF	1	0
PCECC96	Real Personal Consumption Expenditures (Billions of Chained 2005 Dollars)	5	1
UNLPNBS	Nonfarm Business Sector: Unit Nonlabor Payments (Index 2005=100)	5	1
IPDNBS	Nonfarm Business Sector: Implicit Price Deflator (Index 2005=100)	5	1
OUTNFB	Nonfarm Business Sector: Output (Index 2005=100)	5	1
HOANBS	Nonfarm Business Sector: Hours of All Persons (Index 2005=100)	5	1
COMPNFB	Nonfarm Business Sector: Compensation Per Hour (Index 2005=100)	5	1
ULCNFB	Nonfarm Business Sector: Unit Labor Cost (Index 2005=100)	5	1
COMPRNFB	Nonfarm Business Sector: Real Compensation Per Hour (Index 2005=100)	5	1
OPHNFB	Nonfarm Business Sector: Output Per Hour of All Persons (Index 2005=100)	5	1
OPHPBS	Business Sector: Output Per Hour of All Persons (Index 2005=100)	5	1
ULCBS	Business Sector: Unit Labor Cost (Index 2005=100)	5	1
RCPHBS	Business Sector: Real Compensation Per Hour (Index 2005=100)	5	1
HCOMPBS	Business Sector: Compensation Per Hour (Index 2005=100)	5	1
OUTBS	Business Sector: Output (Ineex 2005=100)	5	1
HOABS	Business Sector: Hours of All Persons (Index 2005=100)	5	1
IPDBS	Business Sector: Implicit Price Deflator (Index 2005=100)	5	1
CP	Corporate Profits After Tax	5	0
SP500	S&P 500 Index	5	0